Object Characterisation using Wideband Sonar Pulses

Mariia DMITRIEVA

Ocean Systems Laboratory

School of Engineering and Physical Sciences

Heriot-Watt University

A thesis submitted for the degree of

Doctor of Philosophy

June 2018

Abstract

Characterisation of objects in an underwater environment is challenging. Success in the task can be beneficial in a variety of scenarios, which include oil and gas pipe maintenance, archaeology, and assistance to general underwater object identification. This work focuses on object characterisation, providing a solution for material identification.

To do this, one must sense the underwater environment for which there are several different ways. Some of the most popular rely on sonar images. These provide limited information about the objects, mostly the shape, size and distance to the object. The study of acoustic wave scattering over a wide frequency range provides more information about the targets characteristics. This work builds on the principles of sound scattering. An acoustic echo reflected from an object has a different pulse shape and frequency composition than its initial pulse. These changes in the pulse are due to the interaction of the sound wave with an object during the reflection process and the pulses interaction with the transmission medium. Study of the reflected pulse can provide information about physical properties such as size, material and shell thickness. The objects used in this work are limited to spherical shells made of a variety of materials, and filled with different liquids or air.

The task of material identification is approached in two different ways. The first one is a machine learning based approach. The classification is not based on the object's shape, but on its physical properties including the composition material. Two approaches will be presented: one, where the spherical shell is described by the echo's representation in time-frequency domain and one, where it is described by the form function. The objects are classified using a number of machine learning techniques including support vector machine, gradient boosting and neural networks. The machine learning approaches give good results for a number of tasks, but are not sufficient to distinguish between materials with similar properties, like water and salt water.

An alternative solution is presented in this thesis, which identifies the filler and the shell materials separately. This material identification approach is based on the timing of the sound scattering components. The echo reflected from an object is formed by a number of processes. The information about these processes can be extracted from the echoes and used to identify the material. This approach does not require any training data and shows good results, which are demonstrated on both the simulated and experimental data. This work focuses on object characterisation, providing a solution for material identification using underwater acoustics and signal processing techniques.

Dedication

In loving memory of my mother, Olga Dmitrieva ...

Acknowledgements

I would like to express my sincere gratitude to my supervisors, Prof David Lane, Dr Keith Brown and Prof Gary Heald. I feel extremely privileged to have been their PhD student. I am thankful to David for giving me an opportunity to be part of the Ocean Systems Laboratory and introducing me to the world of underwater robotics. I am very grateful to Keith for always being there and providing me with encouragement and continuous guidance throughout my Ph.D. And I am deeply indebted to Gary for his enthusiasm and interest in my work. He has been a tremendous help during entire period.

I am grateful to all my colleagues in Ocean Systems Laboratory, Robocademy project and to Hydrason Solutions. That was a great period of my life and you, my friends and colleagues, have made this time amazing and unforgettable with all our trials trips, project meetings, participation in robotics competition, nights out together, baking Tuesdays and even a bit of Ceilidh dancing. I would like to mention and thank Nicolas and Yan for their patience answering never-ending questions and for the memorable scooter driving lesson. Especially I would like to thank Corina, with whom I was lucky to share an office and became a good friend. I would also like to thank Bence for his support and invention of the ”disartixia” term.

I am immensely thankful to my family for their love, prayers and care, for staying close even living far away, for making me feel at home through our skype calls. Especially I want to thank my father, whose wisdom and love have made me who I am today. A very special word of thanks goes to Sharib for making me happy in everyday of our life. Thank you for being my inspiration and supporting me in the work on this thesis.

Finally, I gratefully acknowledge the funding for my research study received from the Marie Curie Initial Training Network ROBOCADEMY of the European Union’s Seventh Framework Programme.

HERIOT WATT UNIVERSITY

ACADEMIC REGISTRY
Research Thesis Submission

Name:	Mariia Dmitrieva		
School:	Engineering & Physical Sciences		
Version: *(i.e. First, Resubmission, Final)*	Final	Degree Sought:	PhD

Declaration

In accordance with the appropriate regulations I hereby submit my thesis and I declare that:

1) the thesis embodies the results of my own work and has been composed by myself
2) where appropriate, I have made acknowledgement of the work of others and have made reference to work carried out in collaboration with other persons
3) the thesis is the correct version of the thesis for submission and is the same version as any electronic versions submitted*.
4) my thesis for the award referred to, deposited in the Heriot-Watt University Library, should be made available for loan or photocopying and be available via the Institutional Repository, subject to such conditions as the Librarian may require
5) I understand that as a student of the University I am required to abide by the Regulations of the University and to conform to its discipline.
6) I confirm that the thesis has been verified against plagiarism via an approved plagiarism detection application e.g. Turnitin.

* *Please note that it is the responsibility of the candidate to ensure that the correct version of the thesis is submitted.*

Signature of Candidate:		Date:	

Submission

Submitted By *(name in capitals):*	
Signature of Individual Submitting:	
Date Submitted:	

For Completion in the Student Service Centre (SSC)

Received in the SSC by *(name in capitals):*			
Method of Submission *(Handed in to SSC; posted through internal/external mail):*			
E-thesis Submitted ***(mandatory for final theses)***			
Signature:		Date:	

Please note this form should be bound into the submitted thesis.
Academic Registry/Version (1) August 2016

Contents

List of Figures

List of Tables

List of Symbols

a_{ch}	chirp rate
b_1	lowest frequency of the chirp
λ **and** μ	Lamé constants
u	displacement
ρ	density of the material
k	wave number
r	distance from the source to the object
P_0	amplitude of pressure in incident wave
f_∞	form function
$s_i(t)$	initial pulse
$s_{scattering}(t)$	reflected pulse
a	radius of the sphere
d	thickness of the sphere
$z(t)$	analytical signal
δ	Dirac impulse
K	kernel
r, θ, ϕ	spherical coordinates
Φ	scalar potential
Ψ	vector potential
$C_L^{(n)}$	longitudinal speed of sound in the layer n
$C_L^{(n)}$	transversal speed of sound in the layer n
j_l	Bessel function of the first kind
n_l	Bessel function of the second kind
h_l	Henkel function of the first kind
f_s	sampling frequency
c_{pl}	speed of the plate wave
θ_c	critical angle
Δt_{shell}	timing of the plate wave component
Δt_{filler}	timing of the wave propagating through the filler

List of Abbreviations

AUV	Autonomous Underwater Vehicle
CNN	Convlutional Neural Network
DFT	Discrete Fourier Transform
DTFT	Discrete Time-Fourier transform
FD	Frequency Domain
FF	Form Function
FFT	Fast Fourier Transform
GB	Gradient Boosting
GT	Ground Truth
HF	High Frequency
LF	Low Frequency
MFE	Mid Frequency Enhancement
MLP	Multilayer Perceptron
MPP	Minimum Peak Prominence
ReLU	Rectified Linear Unit
RSPWVD	Reassigned Smoothed Pseudo Wigner-Wille distribution
SNR	Signal-to-Noise Ratio
STFT	Short Time Fourier Transform
SVM	Support Vector Machine
TD	Time Domain
TFD	Time-Frequency Domain
TFR	Time-Frequency Representation
TFM-SVD	Time-Frequency Moment Singular Value Decomposition
WVD	Wigner-Ville Distribution

Chapter 1

Introduction

I am among those who think that science has great beauty. A scientist in his laboratory is not only a technician, he is also a child placed before natural phenomenon, which impress him like a fairy tale.

Marie Curie

1.1 Thesis motivation

The Earth is known as the ”blue planet”, because it is largely covered by water. About 71% of the Earths surface is water. Nowadays the knowledge about the ocean is rapidly increasing, nevertheless the most of the ocean's territory still remains unexplored. Ocean exploration is a wide research area that is in high demand for the oil and gas industry, nature studies, military and archaeology. The task of the ocean exploration can be undertaken using unmanned vehicles, which both include remotely operated vehicles (ROV) and autonomous underwater vehicles (AUV). A ROV is operated by people aboard a vessel, to which it is connected via cables for control and communication. An AUV conducts its mission without operator intervention. The vehicles are equipped with a variety of sensors to assist in given tasks, as well as with control and navigation systems. One of the challenging tasks in the underwater domain is object characterisation, when a detected object has to be described mostly for recognition purposes. An object can be characterised by a number of properties including shape, size, colour, structure and the material that the object is made of. The object characterisation is not a trivial task and requires a proper way of sensing the environment. The choice of the sensor depends on the characteristics of the object, which are required to be obtained.

There are different ways to sense the underwater environment for the task of object characterisation. The electromagnetic waves are commonly used for sensing in the air, but underwater it is not as advantageous. Radio waves are strongly absorbed by sea water and radar (radio detection and ranging) technology cannot be used underwater. Optical sensors, which use visible light spectrum, work at short range, as they strongly affected by

attenuation of light underwater. The quality of the image and the camera range depends strongly on the clarity of the surrounding environment. Additionally, camera images lack the range information. Another type of sensors, which important to mention are magnetic sensors. For example, a magnetometer allows to measure magnetic fields and even though it doesn't provide enough information to identify object, it provides additional description of the target. Acoustic sensors are the most preferable in the underwater, as they use sound waves, which easily propagate underwater over long distances, depending on the frequency used. The underwater acoustic propagation is affected by a number of factors. The acoustic wave expands as a spherical wave in water and the acoustic intensity decreases in inverse proportion to the surface of the sphere. Acoustic absorption is frequency dependent and lower frequencies reach longer than higher frequencies. It is also a function of temperature, salinity and depth (pressure). Noise in underwater environment consists of ambient noise and other interfering noises. Reverberation is unwanted echoes and paths of the transmitted signal. As acoustic waves are the only long range information carrier under water, the sonar is the most common solution for underwater sensing.

Sonar is an acronym from SOund Navigation And Ranging. The technology can be divided into two main groups: active and passive. Active sonar transducers emit an acoustic signal into the water, while passive sonar systems listen without transmitting. Passive sonar detects the noise which can originate from objects, background noise of the environment and other noise sources. In this work only active sonar technology is used. The sonar equation can be used to determine the signal excess of a target, taking into account the level of the background noise (passive) or reverberation (active). The frequency content of the signal can also give clues about the sounds and target type. The transmitted pulse of the active sonar can be at constant frequency or may cover a range of frequencies (wideband sonar). In this work the wideband pulse term is used in a context of a structured chirp-based signal. The standard solution for the object descriptor is using a sonar image. The sonar imaging makes use of acoustic waves to construct a 2D or 3D image of an underwater structures, objects, marine life, sea floor, sub-bottom features and other targets. In this way, the sonar imaging provides information mostly about shape and size of the target. Commonly, a single frequency sonar is used for the imaging. Narrowband sonar can utilize the Doppler effect to detect moving underwater objects. There are also types sonars that do not use images for example, those that use very low frequencies for submarine detection and acoustical oceanography. These sonars arent used for the sea floor analysis and object detection. In addition, there are sonars that are used simply for depth measurements but they are not typically used for sea floor and object analysis.

Wideband sonar with chirp pulses widely used, however not for the study of the wideband echoes, but in the sonar imaging, where chirp pulses allow pulse compression on reception. The pulse compression technique provides improvement in the range resolution as well as the signal-to-noise ratio (SNR) of the processed image [1]. The advanced technique provides high quality of the sonar imaging and used in many applications including pipe inspection, sea-floor survey, also the chirp pulses became a common solution

for fishfinders, which are used to detect fish underwater. Another benefit of the wideband pulses is on the study of acoustic wave scattering over a wide frequency range, which provides more information about the target's characteristics. Chirp pulses have a big advantage over a single frequency pulses, as some of the processes and resonances are happening in objects only at particular frequencies and the wide frequency range of a chirp signal secures their coverage. In nature, pulses with a wide frequency range are used by animals, like dolphins and whales for echolocation. These pulses allow the animals to recognise not only the shape of the object, but its structure and material. Dolphins use double-chirp pulses, which they adapt to the environment and the task [2]. It can be assumed that the double chirp structure of the pulses improves resolution of the received signal, as well as interaction of the chirps between each other and with the target can create extra components, which assist animals in the object recognition task. A number of experiments illustrated that dolphins can be trained to distinguish between plates made of different materials [3]. The dolphin's biosonar has been studied for decades and the research has a big influence into the development of the sonar technologies [4] [5] [6].

The wideband sonar provides additional information to the object's image. It reveals distribution of energy over a wide frequency range and allows the study of a composition in the frequency domain and the timing of the reflection components. The study exposes information about object's characteristics, including object's structure and material, which leads to a better object characterisation. The research presented in this thesis focuses on the characterisation of the object's material by processing wideband pulses. Material recognition can be beneficial for a number of applications, where the general object recognition and classification could be improved by the knowledge about material:

- **Oil and gas industry**: The economic viability, efficiency and safety of the oil and gas industry depends on the transportation of fluids through offshore pipelines. The pipeline construction can be affected by a number of factors including corrosion, changes in marine environment, damages caused by storms. Inspection and repair of the underwater pipelines is essential for spill prevention. Different methods to inspect, repair and maintain underwater pipelines have been developed to ensure that the required fuel can be transported safely, the approaches described in the work will add to the methods available.

- **Underwater archaeology**: This branch of archaeology studies the past through submerged remains, like shipwrecks, submerged landscapes and ruins. The underwater archaeology is challenging in many ways. Underwater sites are difficult to access, they often are dynamic, because of currents or storms, some structures can be buried beneath sediment or occupied by some marine life, it can even be chemically active (for example corrosion). The discipline benefits from remote sensing equipment and use the recorded and analysed data for remains discovery.

- **Power and communications cables inspection**: Most of the Internet data is transmitted by wires built at the bottom of the ocean, called submarine communications

cables. In total, there are hundreds of thousands of miles of communications cables in the ocean floor. A submarine power cables transmit electrical power at high voltage. The cable connection can be disrupted by fishing trawlers, boat anchors, natural disasters, and even shark bites. The maintenance of the submarine communications and power cables would benefit from autonomous tracking and inspection, where the composition of the cable can be identified using wideband technologies.

- **Localisation of the AUVs**: In underwater environment, the localisation is especially challenging due to the lack of the global positioning system (GPS) and a low bandwidth of the underwater communications. The localisation problem is commonly solved by using inertial sensors (doppler velocity logs, compass, pressure sensor) and periodic surfacing of the AUV. Sensing of the surrounding environment and application of simultaneous localization and mapping (SLAM) technology are also an option for the localisation, which would benefit from the better object characterisation.

- **Marine biology**: It is a very large discipline studying life in oceans, which also includes human modification to ocean biology. In 20 century, sonar technology had become essential part of the marine biological research [7]. Different sonar tools had been developed to support different field of study including plankton biology research, fisheries biology, marine mammal research. The study of marine organisms, their behaviors, and their interactions with the environment can benefit from the technology.

Object's shape often not known before the target is detected. Underwater camera or imaging sonar can be used to identify the target's shape. For this work it is important to take into account shape of the object, while working on the material characterisation. The reflected echo is highly affected by the shape and narrowing the possible variety of shapes of the object helps to achieve baseline results. This work is focused on the material recognition for the spherical shells, where two materials are involved - filler and shell.

1.2 Thesis structure

This thesis is structured into six chapters. Chapter 2 introduces a review of the subjects involved into the work. The chapter looks at a number of approaches and considers the relative merit of the research. It covers topics of bio-inspired echolocation, adaptive echolocation, analytical solution for the sound scattering from an object, analysis of the echo's components, signal representation and machine learning approaches.

Chapter 3 demonstrates a solution for the numerical calculation of the sound wave scattering from a spherical shell immersed in water. The solution is adapted from work of Goodman and Stern [8] and Pailhas [9] to the case of 2-layer spherical shell filled with liquid, where the filler material and outside medium are not the same liquids. A set of synthetic

data is generated based on the calculations to investigate the influence of the sphere's properties into the shape and composition of the reflected pulses. It illustrates how the changes in sphere's radius, thickness, filler material and shell material influence the reflected echo.

Chapter 4 presents machine learning solutions for the classification of objects based on their characteristics. Firstly, echoes from a set of spherical shells are classified by a multi-class classifier, where one class is presented by one sphere which can be considered as a combination of characteristics, such as radius, thickness, shell and filler materials. The echo is presented in time-frequency domain (TFD) and fed into a classifier. A number of machine learning techniques are compared for the task, including support vector machine, gradient boosting and neural networks (convolutional neural network and multilayer perceptron). The second presented classifier is based on only one characteristic - filler material. In this way, objects are classified into groups based on their material. The reflected echo is presented by form function and classified using a support vector machine and a multilayer perceptron for comparison. This chapter also illustrates one more application of the machine learning and wideband sonar - sea floor recognition. Different feature extraction methods are applied in combination with machine learning to present classification of the sea floor for data recorded in Tallinn bay, Estonia.

In Chapter 5, the material identification approach is presented for the filler and shell material recognition. This approach is based on the timing of the reflected components and doesn't require any training. Pulse compression, where matched filtering is applied to the reflected echo and initial pulse, and peak detection are used to define timing of the reflection components and calculate speed of sound in the material. The filler and shell material identifiers are presented in separate sections, as the approaches are based on the different reflection components. The results are presented for both the synthetic and experimental data.

Chapter 6 summarises the research and concludes the work presented in the thesis. Suggestions for the future research are outlined in this chapter, presenting possible extension of the thesis.

1.3 Contributions

This research has the following contributions.

- **Extension of the solution for the sound scattering**: In Chapter 3, the solution for the sound wave scattering from a spherical shell immersed in water is extended from work of Goodman and Stern [8] and Pailhas [9] to the case of 2-layer spherical shell filled with liquid, where the filler material and outside medium are not the same liquids.

- **Classification based on the object's characteristics**: Two solutions are presented, based on the time-frequency representation (TFR) and form function (FF) descriptors, using machine learning techniques to classify objects, which have the same

shape, but different characteristics and are filled with different materials. While the first approach classifies objects with a combination of properties, the second approach divides objects into two groups based on their filler material.

- **Novel material identification approach**: The effective approach for the shell and filler material identification is developed in Chapter 5. The approach is based on the timing of the reflection components and doesn't involve any machine learning techniques. The material is identified from the evaluated speed of sound in the material.
- **Adaptive sensing**: The concept of adaptive sensing is presented in Chapter 2, Section 2.2, where the changes in the initial pulse are defined based on the processing of the already received responses. Also adaptation of the pulse is considered in the work and a criteria for the choice of the pulse duration is introduced.

1.4 Publications

The research performed during the work on the topic lead to the following publications:

Published:

1. Mariia Dmitrieva, Matias Valdenegro-Toro, Keith Brown, Gary Heald, David Lane, "Object Classification with Convolution Neural Network based on the Time-Frequency Representation of their echo", Proceeding of the MLSP workshop, Tokyo, Japan, 2017
2. Mariia Dmitrieva, Keith Brown, Gary Heald and David Lane. "Classification of spherical objects based on the Form Function of acoustic echoes", Proceeding of the Underwater Acoustics Conference and Exhibition, Skiathos, Greece, 2017
3. Mariia Dmitrieva, Keith Brown, Gary Heald and David Lane. "Sea floor recognition with bio-inspired echolocation", Proceeding of the OCEANS, Aberdeen, UK, 2017
4. Mariia Dmitrieva, Keith Brown, Gary Heald and David Lane. "Image processing based strategy for adaptive wideband echolocation", ICASSP, Student Forum, New Orleans, USA, 2017
5. Mariia Dmitrieva, Keith Brown, David Lane. "Adaptive bio-inspired signals for better object characterisation", Proceeding of the Living Machine Conference, Barcelona, Spain, 2015

Submitting:

1. Mariia Dmitrieva, Keith Brown, Gary Heald and David Lane. "Material recognition based on the time delay of secondary reflections using wideband chirp pulses". Planning to submit to IET Radar, Sonar & Navigation.

Chapter 2

Background chapter

Look deep into nature, and then you will understand everything better.

Albert Einstein

This chapter presents an overview of the techniques and methods involved in object characterisation using wideband pulses. The subject of research is an intersection of a few disciplines and the literature review includes bio-inspired pulses, physics of sound scattering, signal processing and machine learning. The wideband pulses are inspired by the biosonar of dolphins and whales. Section 2.1 covers the wideband bio-inspired echolocation. In Section 2.2 concept of the adaptive echolocation is presented. Sound scattering from an object, understanding of the components of the echo and its relation to the object's properties are presented in Section 2.3. Section 2.4 describes signal representation with a focus on time-frequency representation. An overview of a machine learning techniques are presented in Section 2.5.

2.1 Bio-inspired wideband echolocation

Dolphins identify objects using their biosonar, which works by emitting short acoustic pulses (clicks) with high bandwidth and high intensity. The echolocation impulses have a double chirp structure. The complex signal structure and frequency range allow animals to collect more information than just a distance to the object [10]. They can evaluate an objects size, shape, material properties and innards of the object. The efficiency of the dolphins clicks has provided inspiration to simulate their signals for echolocation purposes. The animal's echolocation principles have been studied for decades and a lot of information has been obtained about the dolphins clicks, the signals characteristics and their echolocation strategy [11]. Dolphins demonstrate an adaptive control over a set of clicks by changing the interclick interval, frequency components of the signal and other signal parameters. This section covers study of the dolphin's biosonar and bio-inspired pulses.

2.1.1 Dolphin biosonar research

The first steps in the research of animal sonar were made by studying bats' echolocation. The research was made by Spallanzani, when he observed that bats can easily fly in a dark room and published a paper "Letter on the suspicion of a new sense in bats by Abbat Spallanzani" in 1794 [12]. A few years later Jurine discovers that bats lost orientation in the environment, when their ears were plugged with wax. There were no formulated acoustic theory of biosonar at that moment [13]. This came later, 1938, with work published by Pierce and Griffin [14]. They showed that bats emit ultrasound and receive the echoes, and described the acoustic process underlying echolocation. The study of the dolphin biosonar started after the discovery of the bats' echolocation system. In 1947, McBride observed, that dolphins can avoid fine mesh nets even at night [15]. Later, in 1961, Norris with colleagues proved the dolphins' echolocation capability by covering dolphin's eyes. The animals could navigate in a maze using ultrasonic sound [16].

Most of the early study on dolphin biosonar was conducted in the US Navy Marine Mammal Program [11]. The program is summarised in a book written by Ridgway [17]. Within the program Johnson studied hearing in *Tursiop* and found that they could hear sound from $100Hz$ to $150kHz$ [18]. Wood and Evans [19] studied echolocation in toothed whales. The animal's biosonar system was studied in many research centres, including Naval Ocean System Center, Hawaii Institute of Marine Biology in USA, the University of Aarhus in Denmark, the Institute of Ecology and Evolution in Russia, Acoustic Research Laboratory and Marine Mammal Research Laboratory in Singapore [20] [11].

A series of international symposium meetings were dedicated to problems of sensory functions in echolocating animals. In 1991 the symposium was hosted by the USSR Academy of Sciences and the papers of the meeting were published in a book "Marine Mammal Sensory Systems" edited by Kastelein *et al.* [21]. Kastelein in his work specialised in applied acoustic research with marine fauna and contributed a lot in the study of the marine mammal's anatomy, behaviour, acoustics and physiology.

Singurdson [22] studied the dynamics of the dolphin biosonar behaviour during the tasks of search and detection in SPAWAR Systems Center, Bioscience Division, San Diego CA. Multiple object positions on the open ocean field were used to randomized the experiment and provide realistic environment. The experiments were run in two stages: the first, when the object were placed in water with a slightly negative buoyancy, which located the target about 1 m below the surface, and the second, when the test objects were placed on the bottom in an open-ocean field. The dolphins transmitted pulses of $(0.03-0.1)msec$ duration and energy located in a range of $(30-130)kHz$. The peak frequencies are found at $(40-60)kHz$ and $(10-120)kHz$. The animals emit a train of pulses, from a few to hundreds pulses in one train. The study demonstrated importance of the dolphin's previous experience, which influences the current scan pattern, echo-pulse interval and intensity of the pulse.

Nowadays the animals' sonar has been studied in details and it can be seen, that the

sophisticated system allows animals to characterise and detect object, navigate in complex environment and avoid obstacles. Biosonars of bats and dolphins are similar to each other, but evolved in different media, they have some fundamental differences. Bats generate the echolocation sound in the larynx and project it through the vocal tract. The biosonar sounds of different bats can cover a range from 10 kHz to 150 kHz. The signals are frequency-modulated wideband pulses, which contain one or more downward-sweeping harmonics. [10] Both animals perceive the distance to the target from the echo delay.

Dolphins identify objects by emitting short acoustic pulses with high bandwidth and high intensity. They produce the echolocation sound by forcing air through two pairs of lips in the nasal system and transmit the sound through the melon [23]. Their echolocation pulses have a double chirp structure. Dolphins can evaluate object's size, shape, and even internal content, by processing the whole echo from the object [24]. The echolocation clicks are highly directional. The beam emitted in front of the head of the bottlenose dolphin is about 10 - 12 $^\circ$ according to Herzing [25]. The animals change the intensity of the clicks in regard to the surrounding environment. In open water bottlenose dolphins have higher intensity clicks with the peak to peak source levels up to $(210-230)dB$ re 1 μPa [25]. The value is much lower in an enclosed space. Au [3] presented a study of the dolphins' sonar system, where a few tasks were given to dolphins to discover features of their echolocation system such as signal characteristics, transmission and reception of the pulses. Dolphins could distinguish between metallic plates of the same size but different material (aluminium, brass, copper), the same task was repeated for the solid spheres with the same diameter and different material. The animals were managed to recognise different wall thickness. Later Au *et al.* developed a phantom echo technique which was used by Ibsen [26] to determine which part of the echo's bandwidth is used by dolphins for target recognition. Dolphins were asked to discriminate between water filled spheres made of steel and brass. The phantom echo technique allows mimicking different targets by convolving animal's outgoing signal with a transfer function of different targets. The important conclusion of the [26] is that the target discrimination cues of the dolphins were based only on the frequency content.

The adaptive nature of the bottlenose dolphins was described by Au [2]. The measurement of the echolocation signals was made in open water (Kaneohe Bay, Hawaii). The recognition task contained three different targets. Study of the recordings showed that the dolphins change interval between impulses and optimise the signal characteristics to perform the given tasks better. The ability of the animals to adapt echolocation signals, to improve the recognition process, was later demonstrated by adaptation in beluga whale echolocation signals in [27]. The beluga transmitted two clicks in pairs (high frequency pulse is followed by a low frequency one). The object detection tasks were presented to the animals in San Diego Bay, California, and then repeated in the noisier environment - Kaneohe Bay, Hawaii. The ambient noise level at the second destination is typically $(12-17)dB$ greater, than at the first one. The detection tasks involved recognition of a water-filled stainless steel spheres. Results of the test showed the adaptiveness of the bel-

uga's echolocation system. The animals shifted to higher frequencies and greater intensities, when they were placed to the more noisy environment. The shift in frequency range allows to avoid frequency range of the noise, while increase of intensity improves signal-to-noise ratio and compensates for transmitting in higher frequencies. Though it must be considered that there is a possibility that differences in experiments at those two places could also influence the adaptation process. Nevertheless, based on the experiments, it can be concluded that beluga optimises the signal characteristics for the best echolocation task performance, as well as bottlenose dolphins.

A single dolphin's echolocation click is a pulse with a composite structure. The click can be represented as a combination of two chirps with a delay between each other. Capus *et al.* [5] described three phases of the target detection and identification by dolphins: initial search phase, transition phase between search and acquire and the acquire phase. The initial search phase is characterized by wideband signal emission during the search process. When the target is found, the impulses are changed to double chirp structured signals. This is the transition phase between searching and acquire. Variation in the clicks appears during the target investigation process, when dolphin discovers the target to recognise it.

As noted above, dolphins use different clicks. Houser *et al.* [28] studied the dolphin's clicks which were recorded during different recognition tasks. The echolocation clicks were classified into seven groups based on their spectrum. They were labelled according to the number of distinctly bounded regions, number of peaks in the regions, peak frequency, secondary peak frequency (if it exists) and others properties.

Muller *et al.* [29] studied dolphin echolocation clicks and clustered them into four classes based on the center frequency and root mean-square bandwidth. In each class, echoes from specific targets were calculated using acoustic scattering model. The echoes were presented in time-frequency domain with reduced interference distribution (RID) and difference in time-frequency distribution between the targets was evaluated. In the study they showed that modification in the clicks and evaluation of the return echo in time-frequency domain may be fundamental for the target's discrimination by dolphins.

The click's spectrum and other parameters can be changed by the animals from click to click during the target investigation process. It allows dolphins to "tune" the clicks for better object representation. The interval between the clicks does not stay constant during the object recognition task. Dolphins emit one click after receiving and processing the previous one. The animal takes time to process the return echo before sending the next one. That makes the interclick interval longer than two-way-transit (TWT) time, i.e. longer than time needed for sound to reach target and come back. So the interclick interval strongly related to the distance between dolphin and target [2]. Herzing *et al.* [25] specified that the interclick interval grows when the distance to an object increases, as well as the click pressure changes relatively to the distance to the target or to the ambient noise. Also the neural processing time needs to be considered here, the time between receiving a response and emitting a new click, it varies depending on task and animal capabilities. The interclick intervals can vary roughly between 19 and 45 ms [25]. Dolphins can emit series of clicks,

also called trains or packages. The interlick interval in the case is less than TWT time. Animals process the set of clicks together, even though, click parameters are changing inside of the click train. In this way, they adapt outgoing pulse parameters to the task, target and environment.

Dolphin's echolocation was also studied in the Institute of Sound and Vibration Research, University of Southampton. Leighton *et al.* [30] described advantages in usage of nonlinear mathematical functions to process the dolphin-inspired pulses. They noticed a variation in amplitude between clicks in a click sequence emitted by dolphins. The difference can be used to distinguish between targets and bubble clutter. The bubbles in the water create an obstacle for man-made sonar, which can lead to a target loss. There are two proposed techniques: twin inverted pulse sonar (TWIPS) and biased pulse summation sonar (BiaPSS). TWIPS technique relies on the formation of inverted pulses. The technique uses a pair of pulses which are identical except that one is inverted with respect to the other. Summation of the return pulses will enhance non-linear component - bubbles and suppress linear components - the reflection from the target, while subtraction of the pulses highlights target and suppress bubbles. BiaPSS relies on the variation of click amplitude. The technique uses pulses with the same form, while amplitudes of the pulses are scaled down. Summation of the return pulses allows to enhance linear components and partially suppress the non-linear components. Result is a reduced influence of the bubble clutters [31] [6]. In [6], Leighton and White demonstrated sonar and radar schemes, which use TWIPS, BioPSS and TWIPR (twin inverted pulse radar) to detect and distinguish between targets and the clutter.

2.1.2 Bio-inspired pulses

The study of the dolphins clicks inspired a simulation of the signals for echolocation purposes. Based on the work of Houser *et al.* [28] six biosonar signals were synthesized by Capus *et al.* [5]. The impulses have a double chirp structure, formed by down-chirps with different frequency range. The single chirp can be described by Equation 2.1 [5], where a_{ch} is the chirp rate and b_1 is the lowest frequency of the chirp.

$$chirp = cos(2\pi(a_{ch}t^2 + b_1 t)) \tag{2.1}$$

The high frequency chirp has a delay relative to the low frequency chirp. All the six signals have different chirp frequency range and different chirp rate, Table 2.1 [5].

The spectrum of the pulses varies depend on the chirp components of the impulse. It can be changed by the delay between LF and HF chirps in the click, as well as frequency range of both chirps. As a result, there is a variety of spectra and the pulses belong to different groups according to the Houser's classification. For example, signal DC1 (with chirps frequencies $(30-114)kHz$ and $(46-130)kHz$) has an unimodal spectrum, while the DC6 pulse (with chirps frequencies $(30-78)kHz$ and $(82-130)kHz$) has a bimodal spectrum. The echoes from spherical target had been analysed using short-time fractional

Table 2.1. Set of the bio-inspired signals

		Nomainal frequency bands	
Signal	Chirp rate, a ($kHz\ s^{-1}$)	Chirp 1	Chirp 2
DC1	-0.420 x 10^6	30-114 kHz	46-130 kHz
DC2	-0.375 x 10^6	30-105 kHz	55-130 kHz
DC3	-0.330 x 10^6	30-96 kHz	64-130 kHz
DC4	-0.300 x 10^6	30-90 kHz	70-130 kHz
DC5	-0.270 x 10^6	30-84 kHz	76-130 kHz
DC6	-0.240 x 10^6	30-78 kHz	82-130 kHz

Fourier transform. The work presented discriminatory potential of the pulses and significant difference had been demonstrated between pulses. [5]

2.2 Adaptive sonar

The adaptive nature of dolphin's pulses was studied and described by Au [2], Herzing *et al.* [25], Houser *et al.* [28] and others. It became an inspiration for adaptive echolocation. The idea of adaptive echolocation is based on the concept of changes in the emitted pulse to adapt it for the current environment, object and task. The idea of adaptive signalling was mentioned in [28] and [5] as a possible evolution of the sonar, but was not presented as a complete concept.

Pailhas *et al.* [32] classified seven different objects using two wideband pulses: DC1 and DC6 from Table 2.1. The objects contained cone, cone 8° tilted, pipe, tube, rocket head, brick (long), brick (large). The classification was based on the location of the notches in the spectra of their wideband echoes and it was implemented using Euclidian distance metric. It is shown that DC6 pulse performs better in classification of pipes and tubes, while DC1 performs better in classification of cones and discrimination of their orientation. The research in [32] was focused on the object classification using location of the notches and the distance metric, but moreover it illustrated that the choice of the initial pulse influences the classification result.

Qian *et al.* in [33] used three different wideband double-chirp pulses for object classification and compared the results. The objects were three copper cylindrical shells which have a diameter 10*cm* and 40*cm* with different thickness. The authors presented reflected echoes in time-frequency domain with Wigner-Ville distribution and used singular value decomposition to extract target features. The targets were classified with support vector machine classifier. The three pulses showed different results for the same classification task. In this way, the work illustrated that different pulse types can match different targets for optimal recognition and highlighted possible benefits from the adaptive echolocation.

Pulse reflected from an object presents an interaction of the initial impulse with the object during the reflection, that is why the parameters of the pulse are important to highlight one or another characteristic of the object. Also different type of objects can demand dif-

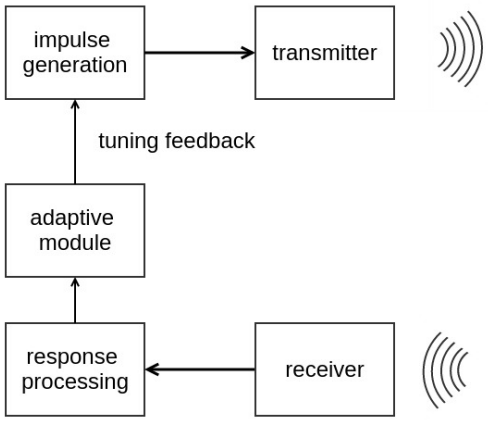


Figure 2.1. Scheme of the proposed adaptive concept

ferent pulses. For example, the spectrum of the echolocation pulse has a big influence on the result. Pulses with different spectrum interact differently with objects. Each object has a resonant frequency or frequencies, so the response will present more information about the object in the case that the click's spectrum excites the resonant frequency.

The adaptive bio-inspired echolocation concept defines changes to be made based on the processing of already received responses. Basic adaptive scheme was present by Dmitrieva *et al.* [34], Figure 2.1. The adaptive module stores previous responses information and tunes the impulse generation module to adapt the next impulse for the particular environment and object. The tuning feedback is formed based on the learned knowledge about current target and environment.

The adaptive strategy first considers distance to the object to set the pulse, then size and shape of the object is taken into account. The final attribute to adapt to is innards of the object (material, structure, thickness of surface, etc). Also the impulse type can be changed according to the task. So the distance to the object and it's shape can be made with one pulse type, while the material recognition is studied by another pulse, which is more suitable for this type of object. The work in [34] presented a conceptual solution for the adaptive approach, but was not implemented on practice.

Attempt to adapt the chirp's slope value based on the representation of the echo structure is presented by Dmitrieva *et al.* [35]. The aim of the approach is to choose a pulse with the best distinction of the dips and peaks in echo structure from a set of pulses with different slope values. The position of the dips and peaks is stable for the different values of the chirp slope, but the depth of the dips and intensity of amplitude of the peaks vary depending on the slope. The concept is based on the evaluation of the peak-dip ratio. A set of pulses with different slopes was processed and pulse with the highest peak/dip ratio in the received echo was chosen for the object characterisation. The approach can be presented with a few steps: (1) time frequency representation (TFR) using smoothed pseudo Wigner Ville distribution, the approach is described in Section 2.4, (2) localisation of the peaks and dips, using image processing techniques and (3) calculation of a peak/dip ratio for the final decision about

the slope value. The paper presents the possible solution for the pulse selection, as a step towards adaptive echolocation.

The adaptive sonar is a relatively new concept. Implementation of this concept requires fine understanding of the relation between the pulse parameters and target's characteristics, environment and task.

2.3 Scattering from an objects

This section provides overview of analytical solutions for the sound scattering by an object submerged in water and a survey on attempts to understand the structure of the echo, as well as its relation to the object's properties.

2.3.1 Calculation of the sound scattering

The sound scattering was first explored mathematically by Rayleigh in 1877 [36]. The proposed solution was limited by the case when the target is small in comparison to the wavelength. Morse [37] derived a solution for the scattering by rigid, immovable circular cylinder and sphere without limit to the target size in comparison to the wavelength. In later publication [38], the solution was extended to include the effect of compressional waves inside the target.

Anderson [39] described reflection from a fluid sphere of a size comparable to the wavelength. The sphere's acoustic properties were close to the properties of the medium and the diameter was up to several wave-length. The author calculated acoustic pressure and the total energy in the scattered wave. In this way, Anderson provided an extension of the Rayleigh work [36], where the sphere's size was comparably small in relation to the wavelength.

One of the earliest investigation of the acoustic scattering from an elastic spherical shell was presented by Junger [40]. He was the first to consider the problem of sound scattering by an air-filled elastic spherical shell and used the thin shell theory for it. In this paper the concept of the separation of scattering solution into a "rigid body scattering" contribution and "radiation scattering" contribution was presented. The concept of separation might still be considered as central to resonance scattering theory [41]. Nevertheless Junger underestimated the extensive effect of fluid loading on the dynamical characteristics of the spherical shell in vacuum and the analysis was regarded as incomplete [42].

Faran [43] presented a work on scattering from a cylinder and sphere of solid materials, which supports shear and compressional waves. The derivation is based on the work of Morse [38] and Love [44]. The numerically computed scattering showed a good match with the experimental measurements. The calculations in [43] presented a basis for the sound scattering calculation and therefore will be presented in more details.

The wave inside a cylinder can be calculated as a solution of the equation of motion of a solid elastic medium, which is describe by Love [44]:

$$(\lambda + 2\mu)\nabla\Delta - \mu\nabla \times (2\tilde{\omega}) = \rho_1 \partial^2 u/\partial t^2, \tag{2.2}$$

where λ and μ are Lamé elastic constants, Δ is dilatation, $2\tilde{\omega}$ is rotation, u - displacement and ρ is density of the target.

Faran provided mathematical expressions for displacement and dilatation inside and for displacement and pressure outside the cylinder and sphere, afterwards applying proper boundary conditions he derived the solution. There are three boundary conditions at the surface of the object: (i) the pressure in the fluid is equal to the normal component of stress in the solid; (ii) the normal component of displacement in the fluid is equal to the normal component of displacement in the solid; (iii) the tangential components of shearing stress disappears at the surface of solid. These conditions were expressed in cylindrical coordinate and spherical coordinates for the cylinder and sphere accordingly. The distribution in angle of pressure in the scattered wave , p_s, at large distance from the object found by Faran are presented in Equation 2.3 for the cylinder and in Equation 2.4 for the sphere [43].

$$|p_s| \xrightarrow[r\to\infty]{P}{}_0 \left(\frac{2}{\pi k_3 r}\right)^{\frac{1}{2}} \left| \sum_{n=0}^{\infty} \varepsilon_n \sin\eta_n exp(j\eta_n)\cos n\theta \right| \tag{2.3}$$

$$|p_s| \xrightarrow[r\to\infty]{} \frac{P_0}{k_3 r} \left| \sum_{n=0}^{\infty} (2n+1)\sin\eta_n exp(j\eta_n) P_n(\cos\theta) \right|, \tag{2.4}$$

where k_3 is the wave number in the fluid surrounded the scatterer, n - order integer, ε_n - Neumann factor ($\eta_0 = 1$, $\eta_n = 2, n > 0$), η_n - scattering phase-angle, P_0 -amplitude of pressure in incident wave, $P_n(\cos\theta)$ - Legendre polynomial.

Hickling [45] published a work which is widely used for obtaining solution for scattering field from a spherical shells. In this work he provided a solution for the case of a homogeneous elastic sphere submerged in water based on the work of Faran [43]. Hickling highlighted a small calculation error in Faran work (misplace coefficients) and derived a solution for a solid elastic sphere in water. He also compared the numerically calculated sound scattering with experimental results.

The peaks and notches on the echo representation in frequency domain, which were shown by Faran [43] were studied by Hickling. He proposed that the changes in echo is caused by vibrations occurring within the solid material of an object and validated the hypothesis for the case of homogeneous solid sphere. It was a very important contribution that Hickling showed that echo's form could provide information about the size and structure of the target. He showed on an example of the homogeneous metallic sphere that it is possible to define its radius by studying peaks and notches of the calculated form function. The form function expresses a pressure field scattered from a target in a range of frequencies, detailed description of the form function is presented in Chapter 3. Hickling showed that there is a significant difference between the form functions of objects with different shapes. In this work he proposed that it is possible to estimate the sphere's size by varying the initial pulse length as well. Furthermore, he was describing the content of the reflecting echo in a few

parts: the first part is a rigid body reflection from the surface (specular reflection) and the following parts of the echo which are reflected by the vibration of the sphere (secondary reflections). He showed it with an example of a 5-cycle pulse. In discussion Hickling divided solid materials into two groups: metallic flint-like substances and substances which are "pliant". The work of Hickling had a big impact on the study of the sound scattering from a spherical shells in water.

Goodman and Stern [8] generalised the solution for the elastic spherical shell. They considered a case when the fluid inside and outside of the shell are identical and obtain expressions for the sound fields inside and outside of the sphere. Several limiting cases were presented in their work: solid sphere, solid elastic sphere in the long wavelength limit, fluid sphere and hemisphere. The results were compared to the work of Faran [43] and Anderson [39].

Doolittle and Uberall [46] obtained a general solution for the scattering field from an infinite elastic circular-cylindrical concentric shell which is submerged in a fluid and contains another fluid. In the same way as Goodman and Stern [8] for the spherical shell, the authors in [46] generalised the results for the cylindrical shell, however considering different liquids inside and outside the shell. Their boundary conditions matching approach is based on the work of Faran [43]. Authors presented number of special cases: the limits of solid elastic cylinder, a fluid cylinder, a thin cylinder with identical and different liquids inside and outside, as well as limits for small solid-elastic, fluid and rigid cylinders.

Pailhas *et al.* [47] extended the results of Goodman and Stern [8]. The authors presented the full analytical solution to elastic multi-layer co-centric sphere. Two possible interfaces between layers are considered: fluid/fluid and fluid/solid. The derived solution for the echo in the time domain of the multi-layer sphere is calculated by Equation 2.5.

$$s_{scattering}(t) = FT^{-1}[\frac{e^{2jk_L^{(1)}r}}{|k_L^{(1)}|r^2}]FT[s_i(t)]f_\infty, \tag{2.5}$$

where FT is Fourier transform, $k_L^{(1)}$ is a wave number in the outer medium and r is a distance from the source to the sphere, s_i is an initial pulse and f_∞ is the form function. The form function can be expressed based on the coefficient by Equation 2.6.

$$f_\infty = \sum_{l=0}^{\infty} i^{l+3} A_l^{(1)} \tag{2.6}$$

The coefficient $A_l^{(1)}$ is defined using boundary conditions at the interface of the sphere layers. The multi-layer spheres are used by authors as a artificial landmarks for underwater simultaneous localisation and mapping [47].

The sound scattering was studied widely for the spherical and cylindrical targets. However, the more realistic objects are usually have a more complex geometry, often having facets, edges, and corners. The sound scattering calculation by an object with a complex geometry is a very complex task. Chinnery *et al.* [48] considered acoustic resonances within

a stadium-shaped cavity and in the following publications [49], [50] described acoustic scattering by one relatively simple faceted geometrical shape - a cube. The form function depends on the angle that the cube is located relative to the receiver and transmitter. In [49] scattering from metallic and glass cube are compared for three cube positions. There are three popular methods capable of solving the scattering problem: the T-matrix method, the finite element method, and a combined finite element/boundary element method. The combined finite element/boundary element method was used to calculate the form function for the cubes. The cubes were ensonified in range from $20kHz$ to $200kHz$. Practical results are close to the calculated ones. In [50] presented an average form function for a few different positioning of a metallic cube. The sound scattering by a complex shape structures cannot be solved analytically. The solution for these complex shapes are provided using finite-element models [51] or boundary-element methods [52]. Neural networks could be also used to define form function and acoustic scattering. Dariochy *et al.* [53] used multilayer perceptron to predict an acoustic form function of an infinite length stainless steel tube. In general, the neural network based approaches require a large amount of training data.

The presented work is focused on the spherical targets which have a simple geometrical shape with a spherical symmetry and with existed analytical solution for the sound scattering. The following section aims to provide inside on the understanding of the scattering by a simple geometrical shapes, such as sphere and cylinder.

2.3.2 Understanding of the reflected echo

Understanding of the processes which occur during the sound scattering from the object and finding the relation between the echo's characteristics and object's property is essential for the object characterisation. This section introduces an overview of the study.

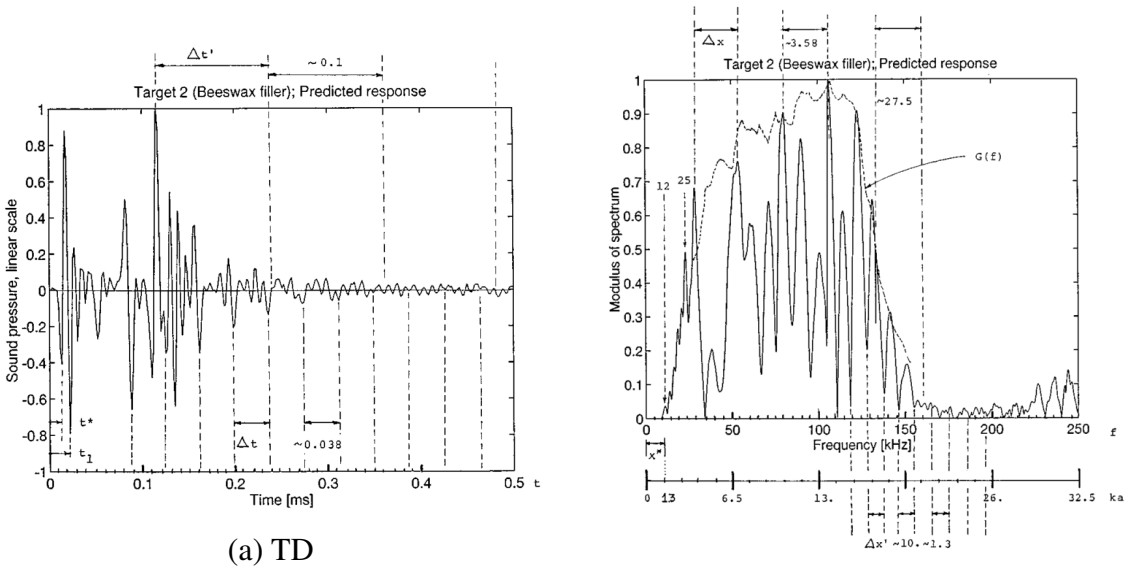
Hickling [54] showed that echo can provide information about the size and structure of a target. He studied peaks and notches of an object's form function and showed that they provide information about the sphere's radius. Kargl and Marston [55] studied a backscattering of a short tone burst from a spherical shell. In this research they presented an idea of using a short duration of the pulses to prevent overlap in time of the Lamb wave with other components. Lamb wave is a surface elastic wave described by Lamb [56]. The incident sound wave (initial pulse) in the water excites a surface elastic wave on the shell. In [55] authors also described the longitudinal wave reverberation, which appears after the specular reflection and proposed that it can be used for the thickness shell evaluation.

Hackman and Sammelman published a series of three papers on the acoustic scattering by a submerged, spherical shell [42] [57] [58]. The first paper [42] was focused on a Lamb wave and was written with the third author Trivet. Two following papers were published in 1991 and focused on other components of the reflected pulse. In the same year Zhang *et al.* published a work on the mid-frequency enhancement of the backscattering of short pulse by thin spherical shells [59]. In the work the authors referred to the research of Hackman *et al.* [42], Kargl and Marston [55], Hickling [45]. This work used pulses with a short

duration to separate in time various echoes occurring in mid-frequency range. The mid-frequency enhancement studied with two approaches. Firstly, the scattering in time domain is evaluated using a discrete Fourier transform and then the amplitude of the mid-frequency scattering is investigated. It is shown, that the mid-frequency enhancement of a scattered pulse has a delay relatively to the specular reflection component and the amplitude of the enhanced scattering exceeds that of the specular reflection. Kaduchak and Marson [60] studied backscattering of chirped bursts by a thick spherical shell. Zhang [59] considered a single frequency initial pulse, while this work [60] investigated chirp pulses. The aim of the investigation was to understand how chirping the frequency of the initial pulse influences the enhancement near the coincident frequency. This work illustrated the enhancement and its dependence on chirp rate for a real shell (the positive rate represents up-chirp pulse and negative - down-chirp pulse). The mid-frequency enhancement of chirped bursts is larger when the frequency changes from high to low, than from low to high (positive chirp rate). Also the authors in [60] suggested that it is possible to estimate thickness of a spherical shell using short chirps. Anderson [61] investigated mid-frequency enhancement of a fluid loaded spherical shells with bistatic measurements. The author used the Wigner-Ville distribution along with acoustic imaging methods for analysis of bistatic scattering. Qiao *et al.* [62] presented modeling and interpretation of the acoustic scattering from stainless steel shell with varying wall thickness using a biomimetic click.

A number of authors have studied different characteristics of the returned echoes and used them to estimates target's parameters. Gaunaurd *et al.* [4] presented processing of an echo reflected by submerged cylindrical shell insonified by dolphin's clicks. Authors presented a predicted and recorded responses from the shell, and described extraction of the target's information from the echoes. The paper describes components of the echo reflected by shell as a combination of the specular reflection from the outer shell's surface, the diffracted circumferential waves evolving around the shell, the elastic waves transmitted into the shell, symmetric and anti-symmetric Lamb waves. The authors highlighted a number of features in the echoes which allow identification of the target's characteristics. There are three features in frequency domain: Δx, $\Delta x'$ and x^* with their equivalents in time domain: Δt, $\Delta t'$ and $t_1 - t^*$. These feature are presented in Figure 2.2 from reference [4].

This approach allows identification of target's radius and thickness, as well as material of the shell and the filler. Δx is found as a distance between significantly large peaks in frequency domain, it is associated with timing Δt. The value of Δx is defined by the ratio of sound speed in shell material and the outer medium. $\Delta t'$ is a timing of the wave travelling back and forth through the filler material, which is represented in frequency domain by $\Delta x'$ - distance between notches in high frequency part of the spectrum. $\Delta t'$ and $\Delta x'$ provide information for the filler material identification. t^* is the arrival of the specular component and t_1 is the arrival of the S_0 Lamb-type wave, which is related to x^* - location of the first resonance peak in frequency domain. Values of $t_1 - t^*$ and x^* depend on the thickness/radius ratio and used by the authors for the thickness evaluation. Shell material is identified using Δx, information about the filler material is provided by $\Delta t'$, while shell's



(a) TD

(b) FD

Figure 2.2. Predicted echo returned by the beeswax-filled shell in time (a) and frequency (b) domains

radius and thickness are evaluated from Δt and x^* accordingly. The measured echo is not identical to the predicted and extraction of the features becoming more complicated. This approach illustrated how dolphins can evaluate the target's parameters without any computations. And the results presented by Muller *et al.* [29] supported the assumption in [4] that dolphins discriminate targets by analysis of the time-frequency information from the returned echo.

Tesei *et al.* [63] proposed an approach which estimates parameters of an air-filled cylindrical shell using resonance scattering analysis. The authors extracted resonance information based on autoregressive (AR) spectral estimation techniques. This technique allows accurate localisation of resonances in the frequency domain. The authors present estimation of the target's outer radius, shell material and shell thickness based on the position of the resonances. The target is an elastic cylindrical thin-walled shell immersed in a salt water. Based on the scattering process there are two resonances which are involved in the parameters calculation: S_0 Lamb-type wave resonance and Scholte-Stoneley wave resonance. The symmetric zero-order S_0 Lamb-type waves are travelling around the shell in the target material. The resonance created by the wave can be detected as distinct and equal spaced dips in frequency domain of the reflected pulse. Scholte-Stoneley waves are travelling around the shell at the shell-fluid interface. Amplitude of the wave reaches maximum near the coincidence frequency, where the Scholte-Stoneley wave travels in phase with the asymmetrical zero-order A_0 Lamb-type wave. The shell's outer radius is calculated based on the Scholte-Stoneley wave, shell material is estimated from the S_0 Lamb wave resonance and the shell thickness from the coincidence frequency. The analysis was presented for a data recorded at sea and provided an accurate target parameter estimation.

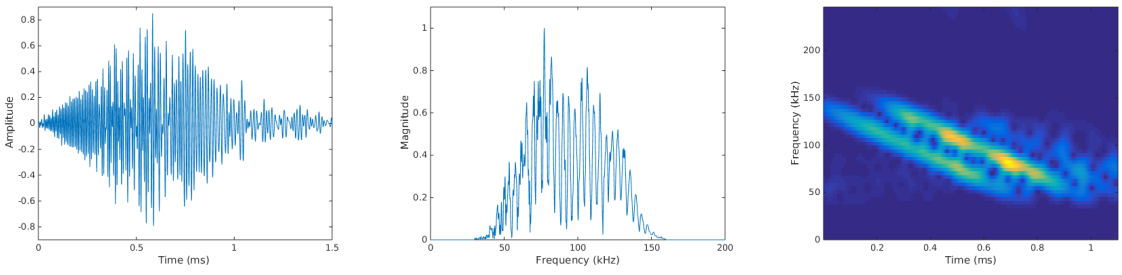
Li *et al.* [41] estimated radius and thickness of a thin spherical shell using mid-frequency enhancement of a short pulse response. The mid-frequency range for the case of reflection

from a thin spherical shell is described as the acoustic wavelength of the range which is small compare to the radius of the shell and large compare to the thickness of the shell: $d \ll \lambda \ll a$. Where d is the thickness and a - radius of the spherical shell. The mid-frequency enhancement occurs next to the coincidence frequency of a thin spherical shell. Estimation of the echo delay and the frequency of the greatest enhancement allowed the authors to estimate radius and thickness of a thin spherical shell. The approach requires prior knowledge about the material of the target, as it influences mid-frequency enhancement characteristics. The presented results are showed the effectiveness of the method.

Pailhas *et al.* [32] analysed and classified wideband echoes reflected from seven different objects based on the location of peaks and dips in the echo's spectra. The authors used bio-inspired synthesised pulses, which have double chirp structure. The objects had different shape, size and were made of different materials. This paper didn't estimate parameters of the targets, but used echo's characteristics for the target's classification.

2.4 Signal representation

To extract important information from a signal and use the information for the effective decision making, it is important to present a signal in a domain, which characterises it the most accurate. The right choice of the domain influences the following step of the pulse's processing. There are two classical signal representations: time domain (TD) or the temporal representation and frequency domain (FD) or the spectral representation [64]. Another way to represent a signal is a time-frequency domain (TFD), which allows to study a signal in both time and frequency domains simultaneously. This work presents analysis of a wideband signals, an example of a wideband echo reflected from an object is presented in Figure 2.3 in TD, FD and TFD.



(a) TD (c) FD (c) TFD

Figure 2.3. Wideband echo with $1.5ms$ duration and $1MHz$ sampling frequency presented in time domain (TD), frequency domain (FD) with 10001 FFT size and time-frequency domain (TFD) using short time Fourier transform with window length $0.15ms$, 1 sample step size and Gaussian window function

The TD representation, $s(t)$ demonstrates how the signal energy is distributed along the time axis. It reveals information about temporal evolution of the signal, its start and end in time, i.e. its actual presence. Working with digital signals, it is important to choose sample rate which allows caption of all the information from the continues signal with

finite bandwidth. The Nyquist-Shannon sampling theorem defines sample rate as $f_s > 2f_{max}$, which excludes aliasing and allows reconstruction of the continues signal. [65] Time domain representation is used in the material identification approach (Chapter 5), where timing of the reflection components is a key component of the material identification.

The Fourier theorem declares that, under some conditions, any bounded signal can be represented as a weighted sum of sines and cosines at different frequencies [66]. The frequency representation of a signal $s(t)$ can be obtained by taking the Fourier transform, Equation 2.7.

$$S(t) = \int_{-\infty}^{\infty} s(t)e^{-j2\pi ft}dt \tag{2.7}$$

Mathematically, the frequency range of a signal is $(-\infty - +\infty)$, but the only positive part is presented on the plots in this thesis. The FD representation shows which frequencies are presented in the signal, minimum and maximum frequency components, relative magnitude of the frequencies, and the bandwidth of the signal. Frequency domain representation is used in Chapter 4, where the difference in the frequency components is used for classification tasks.

TFD representation indicates how the signal energy is distributed along time and frequency axis simultaneously. For this reason, the representation is also called time-frequency distribution (TFD). TFD is specially useful to study wideband pulses as it was presented in [4], [67], [29] and [68]. In this work, difference in TFD of the return echoes is used for object classification (Chapter 4), as it provides image-like representation of the echo and allows to reach high accuracy in classification task. The time-frequency representation of a signal can be achieved by different methods. From historical perspective, Gabor [69] and Ville [70] were the two original papers, which addressed the task of join time-frequency distribution function. Both papers were guided by a development in quantum mechanics [71]. Different time-frequency distributions have been developed since then, and now the solution for the TFR includes spectorgam, short time Fourier transform (STFT), Wigner-Ville distribution (WVD), Page distribution, Choi-Williams distribution (CW), reassigned smoothed pseudo Wigner-Wille distribution (RSPWVD) and many others [64]. This section focuses on the TFD representation and its application for the wideband signals.

Starkhammar and Hansson-Sandsten [68] evaluated different time-frequency representation algorithms to find the best solution for the representation of wideband marine biosonar signals. Authors compare spectrogram, Wigner-Ville distribution (WVD), Choi-Williams distribution (CW), reassigned multitaper spectrogram (RMS) using Hermite functions, locally stationary process multitaper spectrogram (LMS), reassigned smoothed pseudo Wigner-Ville distribution (RSPWVD) and spectrogram multiplication (SM).

The spectogram has several advantages: fast implementation (fast Fourier transform) and straightforward interpretation. However the time slot must be very short to clearly reveal the exact time when the frequency appears [72]. It leads to low frequency resolution and a loss of the difference between closely spaced frequencies. In this way the approach

is a compromise between resolution in time and in frequency. The Wigner-Ville distribution provides the best concentration in time, as well as frequency, but can only be used for signals with a single component in the time-frequency plane. The cross terms appear between all the components in case of more than one components signal and it becomes difficult to interpret the results [73]. Choi-Williams distribution is more smooth in comparison to the previous one, it is also called the exponential distribution [74]. The cross terms are reduced, when the signal components are located at different frequencies and different times. Nevertheless the cross terms remain when time or frequency of two components are the same. Reassigned multitaper spectrogram uses Hermite function as a window [75], which gives the best time-frequency localisation and orthonormality in the time-frequency domain. The advantage is an easy and fast implementation. It smooths the time-frequency representation in a lower time-frequency resolution, which is solved by the reassignment technique. Locally stationary process multitaper spectogram can be considered as a mean square error optimal smoothing kernel and the corresponding multitaper spectrogram representation. Reassigned smoothed pseudo Wigner-Ville distribution is an approach, which first smooths the Wigner distribution and then reassigns components mass to the centre of gravity to recapture the localisation of a single component. Spectogram multiplication can be presented as a multiplication of a long-window and a short-window spectrogram. This operation aims to overcome the disadvantage of the trade-off between a high time resolution and a high frequency resolution. [76]

These time-frequency representation solutions in [68] were applied on three synthetic signals with overlapping in frequency domain, time domain and frequency and time in the same signal, and one recorded click from a bottlenose dolphin. The TFR approaches are analysed based on the way how the two click components clearly localised in both time and frequency domains. Among the presented methods, the reassigned smoothed pseudo Wigner-Ville distribution (RSPWVD) technique was found to be the best one to localize closely spaced signal components.

In context of the comparison and to provide description of the TFR approaches used in this work, short time Fourier transform and Wigner-Ville description will be presented in details.

Short time Fourier transform and spectogram

Short time Fourier transform (STFT) is a Fourier-related transform, where a signal is divided into shorter segments and for each segment a localised spectrum is calculated using Fourier transform. These spectra are plotted as a function of time. The localized spectrum at time $\tau = t$ is calculated by multiplication of the signal $s(\tau)$ with the window $w(\tau)$ centred at the time $\tau = t$ (Equation 2.8), and then computing Fourier transform with respect to τ (Equation 2.9) [64].

$$s_w(t,\tau) = s(\tau)w(\tau - t), \tag{2.8}$$

where $s(\tau)$ is a signal and $w(\tau)$ is a real even window.

$$F_s^w(t,f) = \int_{-\infty}^{+\infty} s(\tau)w(\tau - t)e^{-i2\pi f\tau}d\tau \tag{2.9}$$

The length of the window $w(\tau)$ determines resolution in time and frequency axis. The window provides localisation of the spectrum in time, the shorter the window - the better is resolution in the time domain. On the other hand, the window blurs the frequency spectrum, and the longer window provides higher resolution in frequency domain. In this way, STFT window length is a trade-off between time resolution and frequency resolution. The sensitivity to the window length is a major disadvantage of this approach.

STFT is linear and therefore there is no cross-term artefacts, which are presented in WVD. Simplicity, robustness and freedom from the artefacts are the advantages of the approach.

The spectogram, $S_s^w(t,f)$, is calculated from the STFT. It is the squared magnitude of the STFT expressed by Equation 2.10.

$$S_s^w(t,f) = |F_s^w(t,f)|^2 \tag{2.10}$$

Spectogram carries the same advantages and disadvantages as STFT. The window length problem also leads to a compromise between time resolution and frequency resolution. Even though the spectogram is nonlinear, the nonlinearity is introduced by the power of two only at the last step and it doesn't lead to the cross-term artefacts. The spectogram became a popular method for many applications including acoustics and speech analysis.

Wigner-Ville distribution and its variations

The Wigner-Ville distribution can be described using historical chronology of the development. First, the Wigner distribution, derived in a context of quantum-mechanic [77], Equation 2.11

$$W_s(t,f) = \int_{-\infty}^{\infty} s(t+\frac{\tau}{2})s^*(t-\frac{\tau}{2})e^{-j2\pi f\tau}d\tau \tag{2.11}$$

The WD is a distribution applied to a real signal. The non-linearity of the WD brings interaction between positive-frequency terms and negative-frequency terms. It causes an artefact called cross-term [64]. Ville [70] proposed to use an analytic signal in time-frequency representation instead of a real signal. The analytic signal is a signal with eliminated negative frequency components. It is a complex signal with real and imaginary parts. Usage of the analytic signal before computing Wigner distribution allows the avoidance of the interaction between positive and negative frequency components. As a result, the Wigner-Ville distribution (WVD) is a distribution resulting from an analytic signal being processed through the Wigner distribution, Equation 2.12.

$$W_z(t,f) = \int_{-\infty}^{\infty} z(t+\frac{\tau}{2})z^*(t-\frac{\tau}{2})e^{-j2\pi f\tau}d\tau, \tag{2.12}$$

where $z(t)$ is the analytical signal.

The Wigner-Ville distribution can be only used for simple signals with a single component. Cross terms appear for signals with more then one component in the time-frequency plane. They are placed in between the components. The cross terms are multiplied if the signal has a high level noise. It makes interpretation of a multicomponent signal difficult.

The cross terms can be reduced by smoothing of the distribution with a sliding average window in the time-frequency plane, Equation 2.13.

$$W_{SP}(t,f) = g_1(t) * W_z(t,f) * G_2(f), \tag{2.13}$$

where $g_1(t)$ smoothing window function for time and $G_2(f)$ is a smoothing window function for frequency. The smoothing has a negative effect of spreading out localized signal terms. To rid of the negative effect, the localisation of a single component is recaptured by reassigning its mass to the center of gravity [78].

$$W_{RSP}(t,f) = \int\int_{-\infty}^{\infty} W_{SP}(u,v)\delta(t - t_{SP}(u,v))\delta(f - f_{SP}(u,v))dudv, \tag{2.14}$$

where δ is the Dirac impulse, it relocates the mass of $W_{SP}(u,v)$ to the values specified by calculated values of $t_{sp}(u,v)$ and $f_{sp}(u,v)$ for time and frequency accordingly [68]. This reassignment method allows to improve readability of time-frequency representation. The moving window method was applied to the time-frequency and time-scale representation by Auger and Flandrin [79]. This approach forms a modified version of a TFR by moving its value away from where they are computed, which produces better localisation of the signal components. In this way, the reassigned smoothed pseudo Wigner-Ville distribution (RSPWVD) is an improved version of the Wigner-Ville distribution [80].

2.5 Machine learning approaches for object classification

Machine learning is a subfield of computer science. Over the last few decades machine learning evolved from the study of pattern recognition and computational learning theory in artificial intelligent. Machine learning explores algorithms which learn from data and make a prediction on it. The aim of the machine learning is to understand the data's structure and fit that data into a model, which can be interpreted. It provides solutions for the variety of tasks including classification, recognition, regression, clustering. [81]

All the machine learning methods can be divided into groups: supervised learning, unsupervised learning, semi-supervised learning and reinforcement learning. The supervised learning algorithms are provided with a labelled dataset for training. Unsupervised learning algorithms are trained with unlabelled data. Semi-supervised learning methods use labelled and unlabelled data. Reinforcement learning algorithms use observations and rewards to continuously learn from the environment in an interactive way. A comprehensive review of the machine learning techniques is presented in reference [82]. In this section only su-

pervised learning techniques will be described. The unsupervised techniques do not benefit from already known information (training data), while supervised approaches take advantage of the previous known information, which improves accuracy of classification [83]. The amount of recorded and generated data allows to build a proper dataset and use supervised machine learning in this thesis. Machine learning methods also can be categorised based on the expected output of the algorithm. In this section classification algorithms are described, which are required to support the conducted research.

2.5.1 Support vector machine

Support vector machine (SVM) is a common machine learning technique used for classification and regression. SVM was introduced by Boser *et al.* [84] in context of the Vapnik's statistic learning theory [85] in 1992. They proposed theoretically well founded algorithm which combines two ideas: the use of kernels and the contraction of a hyperplane. The term "support vector" was used for the first time by Cortes and Vapnik later [86].

Support vector machine separates a data with a hyperplane and expands the solution to the non-linear boundaries using kernel [87]. The work of SVM classifier will be considered for a binary classification problem, Figure 2.4(a). The discriminant function for the presented data is non-liner. The data can be mapped into a high dimensional feature space, where it becomes linearly separable, Figure 2.4(b).

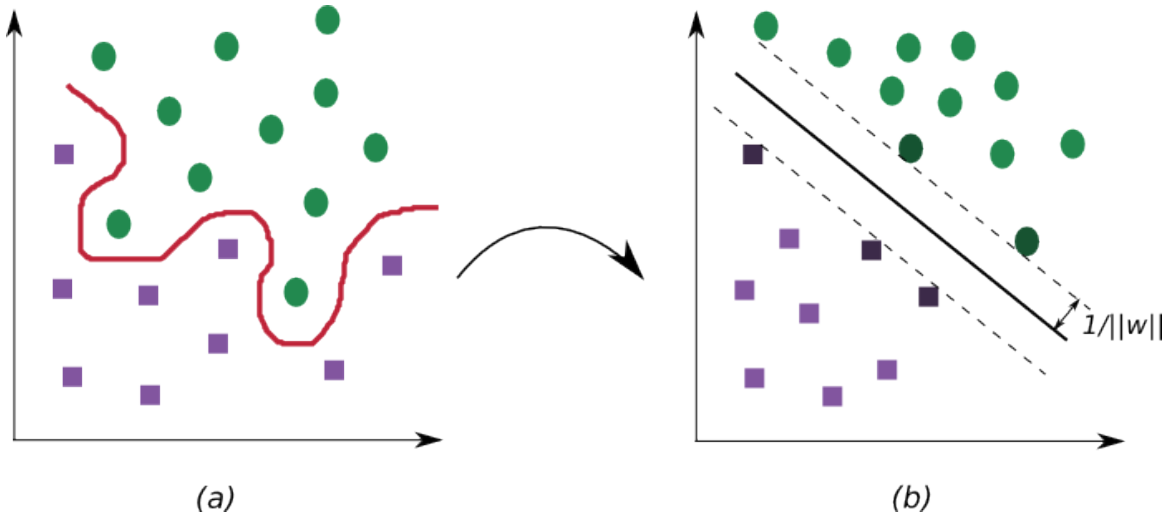


Figure 2.4. SVM: mapping original data (left) into the feature space (right)

The new mapped sample can be presented as $\{(\Phi(\mathbf{x_i}),\mathbf{y_i})\}_{i=1}^{n}$, where for the binary classifier $\mathbf{y_i} \in \{-1,+1\}$. The separating hyperplane is described by equation $w^T\Phi(\mathbf{x})+b=0$. Rescaling w and b for the point which are nearest to the hyperplane in each class, so that $|w^T\Phi(\mathbf{x})+b|=1$. In this way, for every sample $i \in \{1,...,n\}$:

$$w^T\Phi(\mathbf{x_i})+b=\begin{cases}\geq 1, & if\,\mathbf{y_i}=+1\\ \leq -1, & if\,\mathbf{y_i}=-1\end{cases} \tag{2.15}$$

The distance between two classes $2/||w||$ is called margin. An optimal hyperplane is defined as a linear decision function with maximum margin. To define the hyperplane only a small amount of the training data which determine the margin is needed. This training data is called support vectors [86].

The function which characterises the mapping of a data into a high-dimensional feature space is called kernel $K(\mathbf{x},\mathbf{y})$ [88]. The success of the SVM classifier depends on the chosen kernel function. The most common kernel functions are listed below:

linear:

$$K(\mathbf{x},\mathbf{y}) = \mathbf{x}^T\mathbf{y} \tag{2.16}$$

polynomial:

$$K(\mathbf{x},\mathbf{y}) = (c+\mathbf{x}^T\mathbf{y})^d \tag{2.17}$$

Radial Basis Function (RBF):

$$K(\mathbf{x},\mathbf{y}) = exp(-\frac{||\mathbf{x}-\mathbf{y}||^2}{2\sigma^2}) \tag{2.18}$$

SVM has an advantages over the neural networks. It doesn't over generalise, while the neural networks might easily over generalise [87].

2.5.2 Gradient boosting

The gradient boosting algorithm was presented by Friedman [89] in 2001. It belongs to the class of ensemble techniques where a collection of predictors (base learners) are combined together to provide a final prediction. In gradient boosting, the base learners are organised sequentially, to improve with each step based on the loss function. In more details, gradient boosting build additive regression models by fitting a base learner to a current residuals at each iteration sequentially.

The Gradient boosting algorithm is illustrated here for the regression problem with set of variables $\{\mathbf{x} = x_1,...,x_n\}$ and related responses $\{\mathbf{y} = y_1,...,y_n\}$ in the data. The goals is to find an approximation $F(\mathbf{x})$ of the function $F^*(\mathbf{x})$ that maps $\mathbf{x}$ to $\mathbf{y}$:

$$F(\mathbf{x}) = \sum_{m=0}^{n} \beta_m h(\mathbf{x};a_m), \tag{2.19}$$

where $h(\mathbf{x};a_m)$ is a base learner function, β_m is an expansion coefficient. Starting with an initial guess $F_0(\mathbf{x})$, then move to m=1,2,...n, the coefficients for the base learner is updated, Equation 2.20 and the function $F(\mathbf{x})$ is calculated by Equation 2.21.

$$(\beta_m,a_m) = argmin_{\beta,a} \sum_{i-1}^{N} \Psi(\mathbf{y_i},F_{m-s}(\mathbf{x_i}) + \beta h(\mathbf{x_i};a)), \tag{2.20}$$

where $\Psi(\mathbf{y},F(\mathbf{x}))$ is a loss function.

$$F_m(\mathbf{x}) = F_{m-1}(\mathbf{x}) + \beta_m h(\mathbf{x}; a_m) \tag{2.21}$$

The base learner function is an L terminal node regression tree, at each iteration m, a regression tree segments the $\mathbf{x}$ space into L regions $\{R_{lm}\}_1^L$ and predicts a constant value in each of the regions. In this way, the final solution is presented by Equation 2.22 and 2.23 [90].

$$\gamma_{lm} = argmin_\gamma \sum_{\mathbf{x_i} \in R_{lm}} \Psi(\mathbf{y_i}, F_{m-1}(\mathbf{x_i}) + \gamma) \tag{2.22}$$

$$F_m(\mathbf{x}) = F_{m-1}(\mathbf{x}) + \nu\gamma_{lm} 1(\mathbf{x} \in R_{lm}), \tag{2.23}$$

where ν is a parameter to control learning rate, $0 < \nu \leq 1$. The illustrated solution for the regression problem can be easily extended for the classification task, where $\mathbf{y_i}$ is a class label [89]. Gradient boosting is fairly robust to overfitting and a large number of iteration results in better performance [91].

2.5.3 Artificial neural networks

Artificial neural networks are inspired by the structure of the neurons in the brain. Brain consists of a large number of neurons (processing elements), which are connected to each other and form a complex network. It allows brain to carry out complex computations. The neural networks were proposed in the middle of 20th century. The work of McCulloch and Pitts [92] published in 1940s is often recognised as the origin of the neural network [93]. They showed that artificial neural networks can compute any arithmetic or logical functions. Nowadays neural networks achieve cutting-edge performance when computational power and amount of available data have increased dramatically [94]. In this work, neural networks are used for classification, as they are proven to provide high accuracy for the task.

A neural network can be presented as a directed graph. Nodes in the graph correspond to neurons and edges to links between them. Input of a neuron is a weighted sum of the outputs of the connected neurons, Figure 2.5. The activation function simply maps the weighted input into the neuron output. There are several activation functions, the common ones are Sigmoid (Equation 2.24), tanh (Equation 2.25), ReLU (Rectified Linear Unit, Equation 2.26).

$$f(x) = \frac{1}{1+e^{-x}} \tag{2.24}$$

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^x} \tag{2.25}$$

$$f(x) = \max(0, x) \tag{2.26}$$

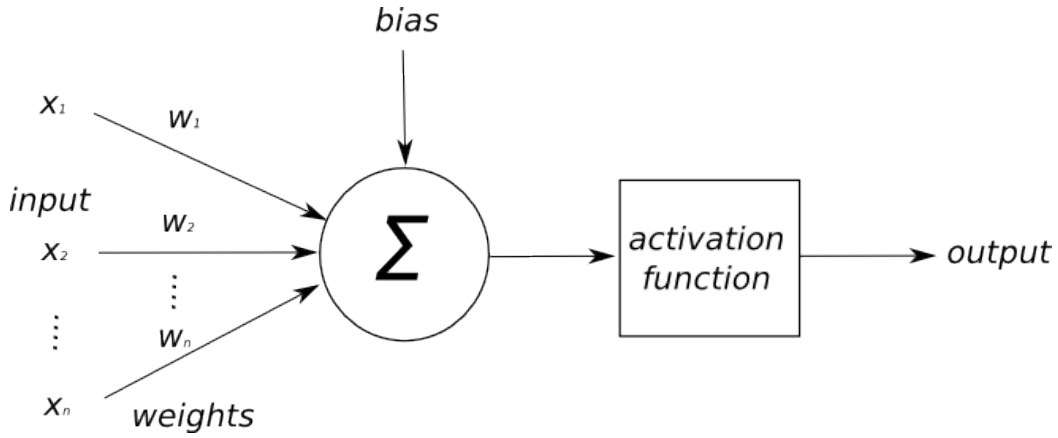


Figure 2.5. Model of a simple neuron in neural network

In this thesis, ReLu activation is used in hidden layers, where it provides faster learning due to the constant gradient. Sigmoid activation is used for the output layer, as it provides well separation of the classes. Neural networks are organised in three interconnected layers: input, hidden and output, Figure 2.6. Multiple hidden layers provide the network with more flexibility and make it more powerful.

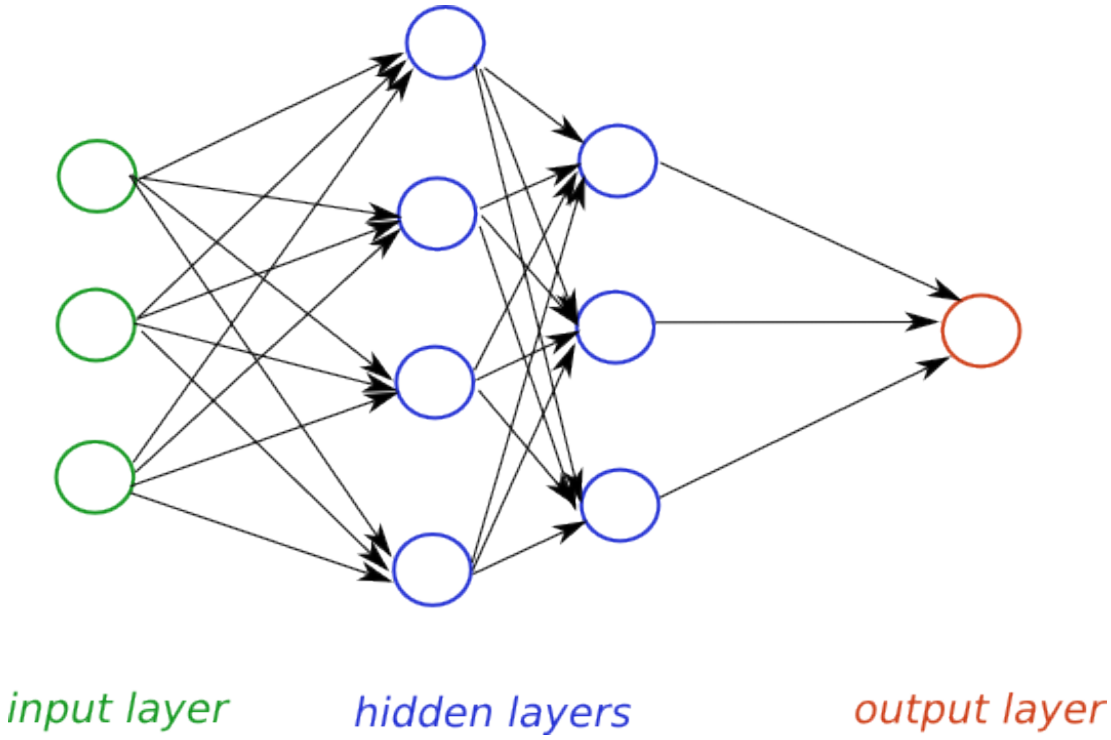


Figure 2.6. Layers of a neural network

The neural networks are trained with back-propagation algorithm [95]. The architecture of the networks can be feedforward, where the connection between layer are only in the forward direction, and recurrent, where there are feedback loops in the network. In this section only feed forwards networks are described.

Multilayer Perceptron

Multilayer perceptron (MLP) is a feedforward network with one or more hidden layers. It consists of fully connected layers (FC). MLP can learn non-linear functions. Figure 2.7 presents an MLP with a single hidden layer. The input layer has three nodes including the bias node and two external inputs.The Hidden layer contains four nodes with one bias. The output layer introduces two outputs of the MLP.

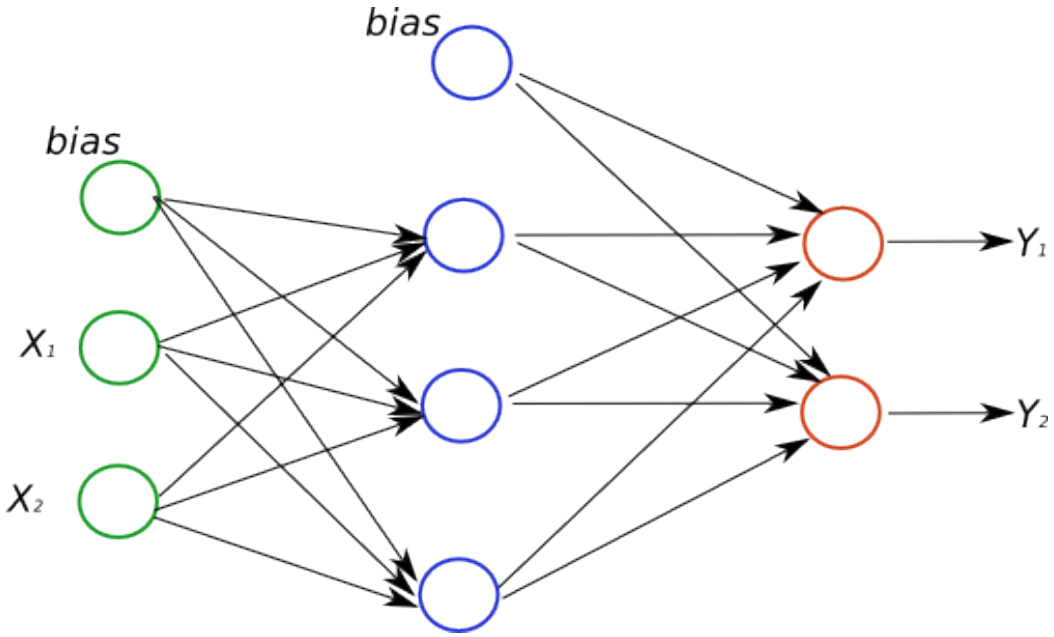


Figure 2.7. Multilayer perceptron with a single hidden layer

The application of the Multilayer perceptron is very wide. It is used in computer vision, pattern recognition, time series prediction, speech recognition, sonar target recognition and others [96].

Convolutional Neural Network

Convolutional neural networks (CNN) consist of one or more convolutional layers. In convolutional layers, the input matrix is convolved with a number of filters, which are sliding though the matrix. The result of this convolution goes to the output of the layer. CNN often contains pooling layers to reduce the size of the representation and speed up calculation. The convolutional layers and pooling layers are followed by fully connected layers. Figure 2.8 illustrates an example of CNN with two convolutional layers, each followed by the pooling layer and two fully connected layers at the end.

The benefit of the CNN is easier training and fewer amount of parameters. CNN are primarily used for image processing and provide results with a high accuracy. Also they are used for the audio signal classification, and there now exists not only 2D, but 1D and 3D CNNs with variety of applications.

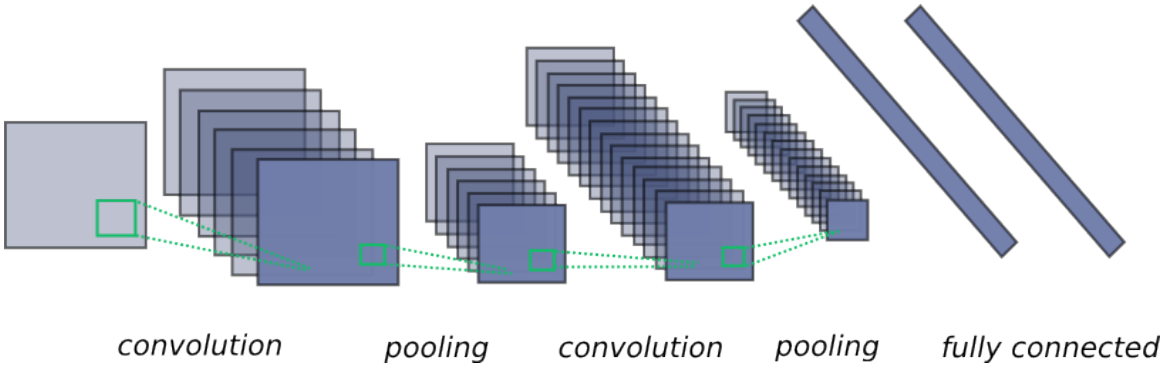


Figure 2.8. Convolutional neural network architecture

2.6 Summary

This chapter presented a literature review on the subject involved in this thesis. It introduced an overview of the techniques and approaches, highlighting those which are the most suitable for this research. Section 2.1 on bio-inspired echolocation described history of the dolphin's biosonar research and bio-inspired pulses. It provided an understanding how the animal's echolocation influenced the wideband sonar development. The bio-inspired pulses described in the chapter have a double-chirp structure. A single linear-chirp pulses are a common solution for the wideband sonar, as they provide the main requirement - covering a wideband of frequencies. The single-chirp pulses provides sufficient information for the purpose of the object characterisation and are exploited in this work. Section 2.2 introduced an adaptive solution for the echolocation task. It presented work which indicated a direction towards the adaptive echolocation, introduced a main concept and a possible solution for the pulse length adaptation. The adaptive approach have a potential to improve object's recognition and classification. This section covered two publication by the author of the thesis [34] [35] on this topic. Section 2.3 presented two important aspects: overview of work on analytical solution for the sound scattering by an object and understanding of the sound scattering components. It introduced research which is exploited in Chapter 3 and Chapter 5. The complete analytical solution for the spherical shell is presented in Chapter 3, while in Chapter 5 some of the sound scattering components are described in details and used for the material recognition. Section 2.4 introduced representation of signals in time, frequency and time-frequency domains. The section is focused on the time-frequency representation, which is used in Chapter 4 in object classification task. Short time Fourier transform is applied in the Chapter 4, because it allows the study of the influence of the time resolution and frequency resolution on the classification accuracy. And the final Section 2.5 introduced machine learning techniques which are used in Chapter 4. From all the variety of the techniques, this section focused on supervised learning, describing the most advanced approaches, which are applicable for the task of classification in this thesis.

Chapter 3

Backscattering of wideband pulses

And indeed all things that are known have number. For it is not possible that anything whatsoever be understood or known without this.
Philolaus

In this work, wideband echoes reflected from an object are processed to collect information about the object's constituent materials. The echoes which cover a wide range of frequencies can be obtained using wideband sonar (Chapter 4 Section 4.1), or numerically calculated based on the knowledge about the object and it's surrounding medium. The wideband coverage can be achieved in a number of ways, for example, by generation of a filtered white noise or by a set of a single frequency pulses. In this work, a linear chirp pulse is used, which gives the advantage of time/frequency separation, hence allowing pulse compression. This chapter presents the numerical calculation of the scattering echo. The approach does not take amplitude value of the initial pulse into account and therefore values of the echo's amplitude are not predicted. The knowledge of the spectrum components variations is enough for the work presented in the thesis.

The target of scattering is limited to a 2-layer spherical shell. The calculated echoes are compared to the recorded wideband echoes scattered from an aluminium sphere. Implementation details are described in Section 3.3. This approach is used to obtain echoes from spheres with different sizes (radius, thickness) and different materials (shell and filler materials). The signals are studied to highlight echo's transformation resulting from the changes in the sphere's characteristics.

3.1 Target description

This work is focused on the scattering from a spherical shell, which is composed of two layers. The outer layer is a solid material and the inner layer is a liquid. The sphere is submerged in a water. The transducer that is used to generate the pulse is placed at a distance r from the sphere. In this way the transmitted pulses propagate in three mediums: water, solid material and filling liquid, Figure 3.1. The sphere shell is made of a solid

material, which supports both longitudinal and transversal waves. Water outside the sphere and a liquid filler propagate only longitudinal wave.

Coordinate system for the sphere is presented in Figure 3.2. The sphere is described using spherical coordinate systems. These coordinates can be expressed in Cartesian coordinates, by Equation 3.1-3.3.

$$r = \sqrt{x^2 + y^2 + z^2} \tag{3.1}$$

$$\theta = tan^{-1}(\frac{y}{x}) \tag{3.2}$$

$$\phi = cos^{-1}(\frac{z}{r}) \tag{3.3}$$

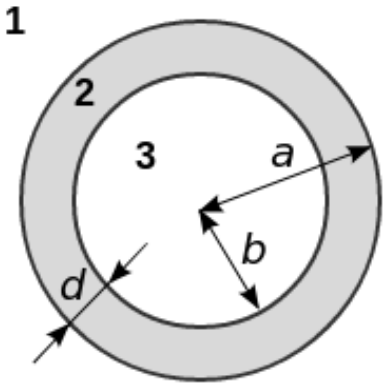


Figure 3.1. Layers of the spherical shell immersed in a water: water (1), solid material (2) and filling liquid (3)

The advantage of using a spherical target is that it axi-symmetrical in all planes, thus not resulting in any angular dependant scattering. The spherical shape of the object gives the same results for any different point of view on the object. Therefore scattering is calculated for one point of view on an object.

3.2 Numerical calculation of the scattering

In this section the scattering from a spherical shell is numerically calculated based on the parameters of the sphere itself and the surrounding medium. This analytical solution is possible only for simple geometrical shapes (such as sphere, cylinder, cube). The calculations are based on the work of Hickling [45], Goodman and Stern [8] and Pailhas [47]. A literature review on the topic is presented in the Chapter 2 Section 2.3.

The scattering is determined in two steps. First the form function of the sphere, f_∞, is calculated, based on the knowledge about shell and fillers materials (layer 2 and 3), water outside the sphere (layer 1) and a size of the sphere (radius and thickness). The scattering echo, $s_{scattering}(t)$, is calculated from the form function and initial pulse, $s_{initial}(t)$, with a knowledge about the distance from the transducer to the sphere.

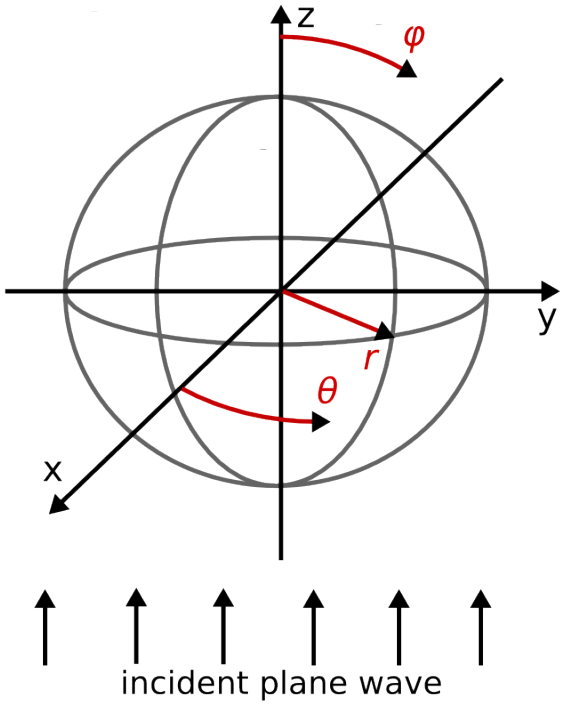


Figure 3.2. Coordinate system of the sphere

3.2.1 Form function

The form function expresses a pressure field scattered from a target in a range of frequencies [97], Figure 3.3. The function is a dimensionless measure of the complex pressure field amplitude which is created by the object during the scattering process [97]. Therefore the form function is a description of an object itself. It depends on the object's size, shape, structure and materials.

A key consideration is that the form function is highly dependent on the view angle, unless the object has a symmetrical spherical shape. In this work we calculate scattering from a spherical object, where the scattering process is not influenced by the view angle.

Form function calculation

Based on the elastic theory, displacement of the wave's energy can be presented by a scalar potential and a vector potential [43], Equation 3.4.

$$u = -\nabla\Phi + \nabla \times \Psi \tag{3.4}$$

The scalar potential Φ associated with longitudinal waves and the vector potential Ψ - with transverse waves. Vector potential component is presented only in solid media, while scalar potential as well as longitudinal wave exists in both solid and liquid media.

The potentials satisfy the equations:

$$\nabla^2\Phi = \frac{1}{(C_L^{(n)})^2}\frac{\partial^2\Phi}{\partial t^2} \tag{3.5}$$

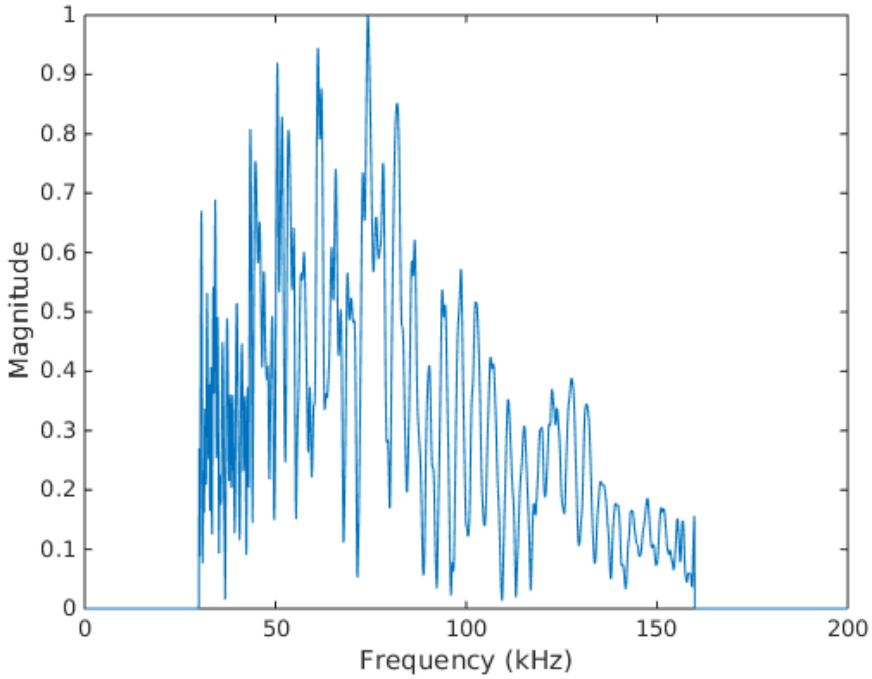


Figure 3.3. Form function of an aluminium sphere filled with fresh water calculated for the range from $30kHz$ to $160kHz$

$$\nabla^2\Psi = \frac{1}{(C_T^{(n)})^2}\frac{\partial^2\Psi}{\partial t^2}, \tag{3.6}$$

where $C_L^{(n)}$ is a longitudinal speed of sound and $C_T^{(n)}$ is a transversal speed of sound in the layer n, Equation 3.7 and 3.8.

$$C_L^{(n)} = \sqrt{\frac{\lambda_n + 2\mu_n}{\rho_n}} \tag{3.7}$$

$$C_T^{(n)} = \sqrt{\frac{\mu_n}{\rho_n}}, \tag{3.8}$$

where λ_n and μ_n are the Lame 1st and 2nd parameters of the layer n, and ρ_n is density of the layer n.

The scalar and vector potential field components of the displacement for each layer can be expressed based on [47]. The potentials below are defined for the particular case of the 2-layer sphere. The scalar potential components for each layer and incoming wave in the outer medium are described in Equations 3.9-3.12.

$$\Phi_{inc} = \sum_{l=0}^{\infty} i^l(2l+1)P_l(cos\theta)j_l(k_L^{(1)}r) \tag{3.9}$$

$$\Phi_1 = \sum_{l=0}^{\infty} P_l(cos\theta)h_l(k_L^{(1)}r) \tag{3.10}$$

$$\Phi_2 = \sum_{l=0}^{\infty} P_l(cos\theta)[A_l^{(2)} j_l(k_L^{(2)} r) + B_l^{(2)} n_l(k_L^{(2)} r)] \tag{3.11}$$

$$\Phi_3 = \sum_{l=0}^{\infty} P_l(cos\theta)[A_l^{(3)} j_l(k_L^{(3)} r)] \tag{3.12}$$

The vector potential exists only in the solid medium, layer 2:

$$\Psi_2 = \sum_{l=0}^{\infty} \frac{\partial P_l(cos\theta)}{\partial \theta}[C_l^{(2)} j_l(k_T^{(2)} r) + D_l^{(2)} n_l(k_T^{(2)} r)] \tag{3.13}$$

In the Equations 3.9 - 3.13 the time-dependent term e^{-iwt} is assumed for all the components [8]. Potentials are calculated for each mode l separately and then summed. The Bessel function of the first kind, j_l, Bessel function of the second kind, n_l, and Henkel function of the first kind, h_l, are all calculated for the spherical case. P_l is Legendre Polynomial, $k_L^{(n)}$ and $k_T^{(n)}$ are wave numbers for the frequency f of the n layer for longitudinal and transversal waves, Equation 3.14.

$$\begin{aligned} k_L^{(n)} &= \frac{2\pi f}{c_L^{(n)}} \\ k_T^{(n)} &= \frac{2\pi f}{c_T^{(n)}} \end{aligned} \tag{3.14}$$

The unknown variables $A_l^{(n)}, B_l^{(n)}, C_l^{(n)}$ and $D_l^{(n)}$ relate to the mode l and layer n. The form function f_∞ can be expressed based on the coefficient $A_l^{(1)}$ [47], Equation 3.15.

$$f_\infty = \sum_{l=0}^{\infty} i^{l+3} A_l^{(1)} \tag{3.15}$$

The coefficient $A_l^{(1)}$ is defined using boundary conditions at the interface of the sphere layers. There are three limit conditions for each fluid-solid/solid-fluid interface [47]:

- The pressure in the fluid must be equal to the normal component of stress in the solid, Equation 3.16

$$p_n^{(n)} = -\lambda_{n+1} \nabla^2 \Phi_{n+1} - 2\mu_{(n+1)} \frac{\partial u_r^{(n+1)}}{\partial r} \tag{3.16}$$

- The normal component of displacement of the fluid must be equal to the normal component of displacement of the solid, Equation 3.17

$$u_r^{(n)} = u_r^{(n+1)} \tag{3.17}$$

- The tangential components of shearing stress must become zero at the surface of the solid, Equation 3.18.

$$\frac{\partial u_\theta^{(n+1))}}{\partial r} - \frac{u_\theta^{(n+1)}}{r} + \frac{1}{r} \frac{\partial u_r^{(n+1)}}{\partial \theta} = 0 \tag{3.18}$$

There are two boundaries for the case of the two layer sphere submerged into water: water (layer 1) - solid (layer 2) and solid (layer 2) - liquid (layer 3). Therefore six equations can be united into one system to find coefficient $A_l^{(1)}$, Equation 3.19.

$$\begin{cases} p_i^{(1)} + p_s^{(1)} = -\lambda_2 \nabla^2 \Phi_2 - 2\mu_2 \frac{\partial u_r^2}{\partial r} \\ u_r^{(1)} + u_s^{(1)} = u_r^{(2)} \\ \frac{\partial u_\theta^{(2)}}{\partial r} - \frac{u_\theta^{(2)}}{r} + \frac{1}{r}\frac{\partial u_r^{(2)}}{\partial \theta} = 0 \\ p^{(3)} = -\lambda_2 \nabla^2 \Phi_2 - 2\mu_2 \frac{\partial u_r^{(2)}}{\partial r_2} \\ u_r^{(3)} = u_r^{(2)} \\ \frac{\partial u_\theta^{(2)}}{\partial r_2} - \frac{u_\theta^{(2)}}{r_2} + \frac{1}{r_2}\frac{\partial u_r^{(2)}}{\partial \theta} = 0 \end{cases} \tag{3.19}$$

Pressure and displacement components can be expressed in terms of scalar and vector potentials:

$$p_n = \omega^2 \rho_n \Phi_n \tag{3.20}$$

$$u_r = -\frac{\partial \Phi}{\partial r} + \frac{1}{r}\frac{\partial \Psi}{\partial \theta} + \frac{1}{r}\frac{cos\theta}{sin\theta}\Psi \tag{3.21}$$

$$u_\theta = -\frac{1}{r}\frac{\partial \Phi}{\partial \theta} - \frac{1}{r}\Psi - \frac{\partial \Psi}{\partial r} \tag{3.22}$$

After some calculations to simplify the expressions, the system 3.24 can be formulated in terms of the coefficient $A_l^{(1)}, A_l^{(2)}, B_l^{(2)}, C_l^{(2)}, D_l^{(2)}, A_l^{(3)}$, Equation 3.24.

The terms x_n and y_n are defined by the Equation 3.23.

$$\begin{aligned} x_1 &= k_L^{(1)} a \\ x_L &= k_L^{(2)} a \\ y_L &= k_L^{(2)} b \\ x_T &= k_T^{(2)} a \\ y_T &= k_T^{(2)} b \\ y_3 &= k_L^{(3)} b \end{aligned} \tag{3.23}$$

$$
\begin{cases}
(A^{(1)}\frac{\rho_1}{\rho_2}h_l(x_1)+A^{(2)}\frac{\lambda_2 j_l(x_L)-2\mu_2 j_l''(x_L)}{(\lambda_2+2\mu_2)}+B^{(2)}\frac{\lambda_2 n_l(x_L)-2\mu_2 n_l''(x_L)}{(\lambda_2+2\mu_2)}-\ldots \\
-C^{(2)}\frac{2l(l+1)}{x_T^2}(x_T j_l'(x_T)-j_l(x_T))-D^{(2)}\frac{2l(l+1)}{x_T^2}(x_T n_l'(x_T)-n_l(x_T)))(-i)^l(2l+1)=\ldots \\
=\frac{p_1}{p_2}j_l(x_1)(-i)^l(2l+1) \\
\\
(A^{(1)}x_1 h_l'(x_1)+A^{(2)}x_L j_l'(x_L)+B^{(2)}x_L n_l'(x_L)+C^{(2)}l(l+1)j_l(x_T)+\ldots \\
+D^{(2)}l(l+1)n_l(x_T))(-i)^l(2l+1)=x_1 j_l'(x_1)(-i)^l(2l+1) \\
\\
(A^{(2)}2(x_L j_l'(x_L)-j_l(x_L))+B^{(2)}2(x_L n_l'(x_L)-n_l(x_L))+C^{(2)}(x_T^2 j_l''(x_T)+\ldots \\
+(l+2)(l-1)j_l(x_T))+D^{(2)}(x_T^2 n_l''(x_T)+(l+2)(l-1)n_l(x_T)))(-i)^l(2l+1)=0 \\
\\
(A^{(2)}\frac{\lambda_2 j_l(y_L)-2\mu_2 j_l''(y_L)}{(\lambda_2+2\mu_2)}+B^{(2)}\frac{\lambda_2 n_l(y_L)-2\mu_2 n_l''(y_L)}{(\lambda_2+2\mu_2)}--C^{(2)}\frac{2l(l+1)}{y_T^2}(y_T j_l'(y_T)-j_l(y_T))-\ldots \\
-D^{(2)}\frac{2l(l+1)}{y_T^2}(y_T n_l'(y_T)-n_l(y_T))+A^{(3)}\frac{\rho_3}{\rho_2}h_l(y_3))(-i)^l(2l+1)=0 \\
\\
(A^{(2)}y_L j_l'(y_L)+B^{(2)}y_L n_l'(y_L)+C^{(2)}l(l+1)j_l(y_T)+\ldots \\
+D^{(2)}l(l+1)n_l(y_T)+A^{(3)}y_3 h_l'(y_3))(-i)^l(2l+1)=0 \\
\\
(A^{(2)}2(y_L j_l'(y_L)-j_l(y_L))+B^{(2)}2(y_L n_l'(y_L)-n_l(y_L))+C^{(2)}(y_T^2 j_l''(y_T)+\ldots \\
+(l+2)(l-1)j_l(y_T))+D^{(2)}(y_T^2 n_l''(y_T)+(l+2)(l-1)n_l(y_T)))(-i)^l(2l+1)=0
\end{cases}
\tag{3.24}
$$

Coefficient $A_l^{(1)}$ can be found from the system of equations 3.24 using Cramer's method [8], Equation 3.25.

$$
A_l^{(1)}=\frac{det(M)}{det(M_{A_l^{(1)}})}, \tag{3.25}
$$

where M is the Cramer's coefficient matrix from the Equation 3.24 and $M_{A_l^{(1)}}$ is the coefficient matrix with $A_l^{(1)}$ column replaced by the right-hand side of the system (the answer values).

The calculated form function is truncated by the number of modes in the sum, equation 3.15. The minimum number of modes N_l is correlated to a sphere radius a and a wave number k. It is restricted by $N_l > 2ka/2$ [32]. The condition for the maximum number of modes is defined by Bessel functions of the first and the second kind, which are used in computations. The functions are bent to zero at the beginning of the frequency axis, while the mode number is increasing. Therefore the maximum number of modes l_{max} is constrained by the desired lowest frequency of the form function.

3.2.2 Scattering

Scattering from the object can be calculated based on the form function f_∞ and initial pulse $s_{initial}(t)$, Equation 3.26 [47].

$$s_{scattering}(t) = FT^{-1}[\frac{e^{2jk_L^{(1)}r}}{|k_L^{(1)}|r^2}]FT[s_{initial}(t)]f_\infty, \tag{3.26}$$

where FT is Fourier transform, FT^{-1} is an inverse Fourier transform, $k_L^{(1)}$ is a wave number in the outer medium and r is a distance from the source to the sphere. The term $e^{-2jk_L^{(1)}r}$ represents a delay based on the distance to the target, while $\frac{1}{|k_L^{(1)}|r^2}$ represents a dispersion during the wave propagation [54].

3.3 Implementation

The described solution for the calculation of the scattering echo can be performed by Algorithm 1. Input contains sphere size (radius a and thickness d), distance from the source to the object, r, parameters of the sphere's materials and outer medium - density $\rho^{(n)}$, speed of sound $C_L^{(n)}$ and $C_T^{(n)}$, Lame parameters $\mu^{(n)}$ and $\lambda^{(n)}$, maximum number of the modes, l_{max}, sampling frequency f_s, frequency range for the form function f_{low}, f_{high} and initial pulse $s_{initial}(t)$. Lame parameters are calculated once for the material, while the wave numbers $k_L^{(n)}$ and $k_T^{(n)}$ are calculated for each frequency separately, as well as x_n and y_n.

Coefficient $A_l^{(1)}$ is calculated for a set of frequencies in a range $[f_{low} : f_{high}]$ (inner for loop). This array $A_l(f_{low} : f_{high})$ and coefficient k_l are defined for each mode l separately (outer for loop). The maximum number of modes is less then l_{max}. It is limited by the values of the Bessel functions for the lowest frequency of the range. The limitation is expressed by the if condition. When the Bessel functions are bent to zero at f_{low}, that coefficient $A_l^{(1)}$ cannot be defined and became NaN . At this point the outer for loop breaks. This way the form function is calculated for the maximum number of modes acceptable for the given sphere and frequency range.

The initial pulse, $s_{initial}(t)$ is presented in a form of analytic signal, Equation 3.27

$$S_{initial}(f) = 0 \quad for \quad f < 0 \tag{3.27}$$

The analytic initial pulse is used for the scattering calculation. The calculated echo is an analytic signal, as well. It contains only positive frequency components. The real signal can be obtained from the analytic by taking only real part of the signal after inverse Fourier transform (IFT) [64].

Algorithm 1 Scattering calculation

Input r, a, d, $\rho^{(1)}$, $\rho^{(2)}$,$\rho^{(3)}$, $C_L^{(1)}$, $C_L^{(2)}$,$C_L^{(3)}$, $C_T^{(2)}$,l_{max}, f_s, f_{low}, f_{high}, $s_{initial}(t)$

Output $s_{scattering}(t)$

calculate Lame parameter $\lambda_2 = \rho_2((C_L^{(2)})^2 - 2(C_T^{(2)})^2)$; $\mu_2 = \rho_2(C_T^{(2)})^2$
$S_{initial}(f) = FT(s_{initial}(t))$
$S_{initial}(f) \rightarrow S_{initial}^{analytic}(f)$
for $l = 0 : l_{max}$ **do**
    for $frequency = f_{low} : f_{high}$ **do**
        calculate $k_L^{(1)}, k_L^{(2)}, ..., x_1, y_3$
        calculate M_l, $M_{A_l^{(1)}}$
        find $A_l^{(1)}$: $A_l^{(1)} = \frac{det(M_l)}{det(M_{A_l^{(1)}})}$
        calculate the coefficient value: $k_l = i^{l+3}$
    end
    if $isNaN(A_l^{(1)}(f_{low}))$ **then**
        break
    end
    save $A_l(f_{low} : f_{high})$ and k_l
end
$f_\infty = \sum_{m=0}^{l_{max}}(A_l^{(1)} k_l)$
$S_{scattering}^{analytic}(f) = \frac{e^{2jk_L^{(1)}r}}{|k_L^{(1)}|r^2}] S_{initial}^{analytic}(f) f_\infty$
$s_{scattering}(t) = real(IFT(S_{scattering}^{analytic})$

This solution was implemented in Matlab. An example of the numerically calculated scattering is presented in Figure 3.4. It demonstrates a $0.1ms$ chirp pulse generated in frequency range $30kHz - 160kHz$ with the 3 dB bandwidth of $(136 - 52)kHz$ reflected from an aluminium sphere filled with fresh water.

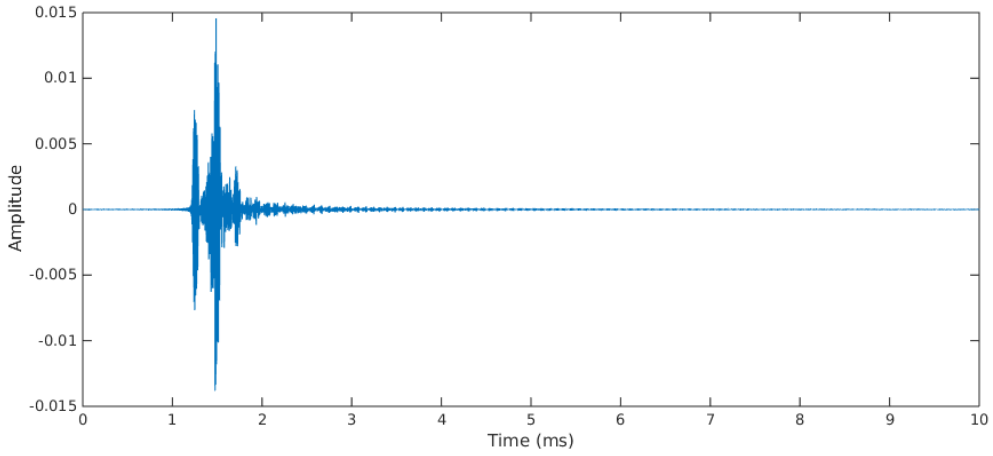


Figure 3.4. Scattering from an aluminium sphere filled with fresh water

The usage of the synthetic data allows to work with a bigger range of objects and not

being limited by available physical objects. Also it provides the possibility to study the influence of the sphere's characteristics into the form of reflected pulses. To rely on the obtained data, the numerically calculated pulse is compared to the recorded pulse.

3.4 Comparison with experimental data

In order to validate the approach, reflected echo is calculated numerically for characteristics of a real physical sphere. The sphere was insonified with the same pulse and the calculated reflection is compared to the recorded experimental data. In this section the results are compared to the data recorded in the test tank of the Ocean Systems laboratory. The set up of the experiment is described in the Chapter 4 Section 4.1. The object is an aluminium sphere, radius $0.075m$ and shell thickness $0.001m$, Figure 3.5.

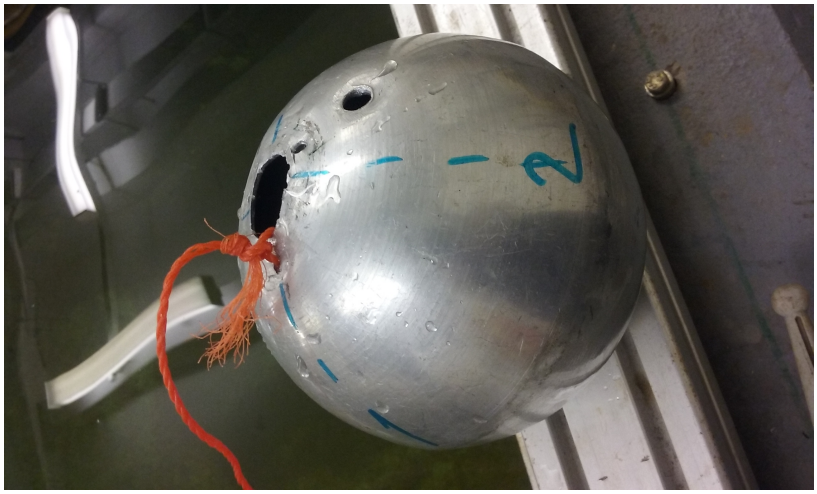


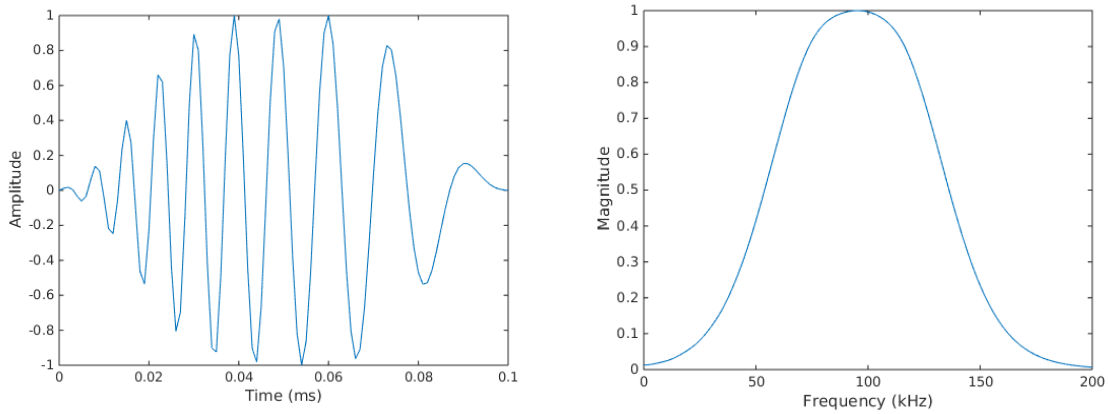
Figure 3.5. Image of the taget: aluminium sphere with radius $0.075m$ and shell thickness $0.001m$

The sphere was filled with water and placed in the water tank. It was filled with the liquid manually and the best attempt was made to leave no air after the sphere was sealed. Parameters of the layers are presented in Table 3.1.

Table 3.1. Parameters of the target's materials

N	material	ρ (kg/m^3)	C_L (m/s)	C_T (m/s)
1	water	1000	1480	-
2	aluminium	2700	6320	3130
3	water	1000	1480	-

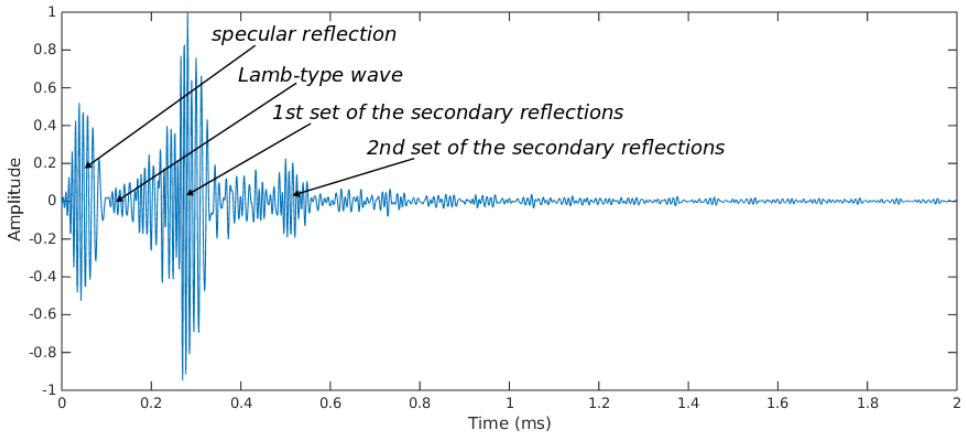
The sphere was insonified with a single down-chirp pulse. The pulse is generated in a frequency range $(160-30)kHz$ with the 3 dB bandwidth of $(136-52)kHz$, duration $0.1ms$ and Blackman windowing, Figure 3.6. The short signal allows allocation of specular and secondary reflections separately. It is beneficial for the comparison, because more details about the scattering process can be observed. Figure 3.7 presents first $2ms$ of the scattering. The analytically calculated data is computed using the method described in this chapter.



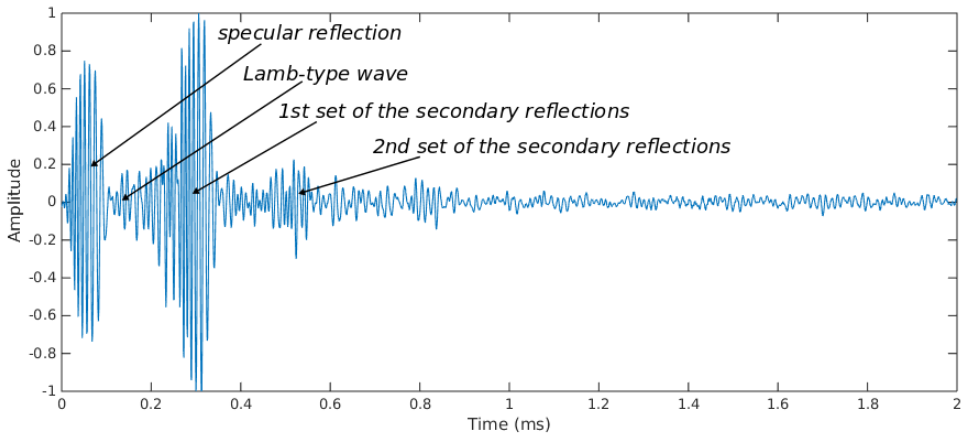
(a) TD (c) FD

Figure 3.6. Generated pulse - linear down-chirp with the 3 dB bandwidth of $(136-52)kHz$, duration $0.1ms$, sampling frequency $1MHz$ and 10001 FFT size.

Experimental data is recorded in test tank, the experimental set-up and sonar system are described in the Chapter 4 Section 4.1.



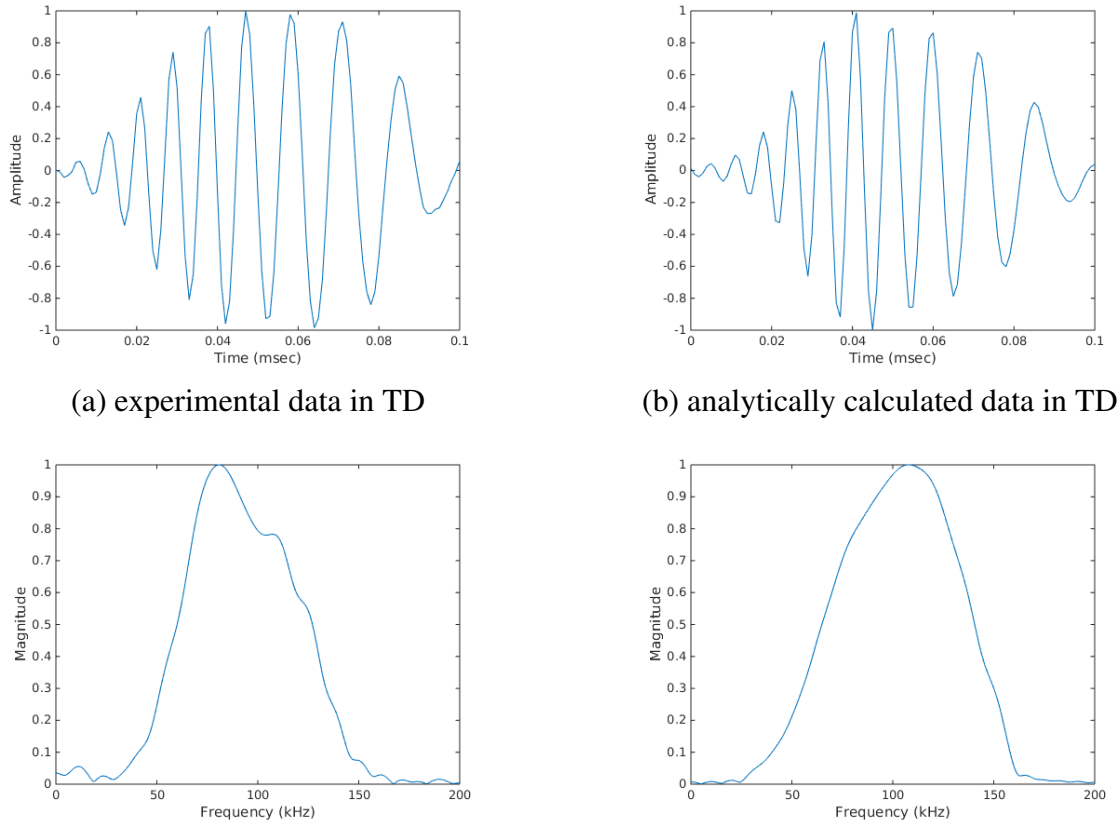
(a) Analytically calculated data



(b) Experimental data

Figure 3.7. Comparison of the analytically calculated and experimental echoes reflected from an aluminium sphere represented in time domain: recording duration $2ms$, sampling frequency $1MHz$

The presented echoes can be studied based on the positioning of their reflection components. The first component is a specular reflection, which has a duration of the initial pulse $0.1ms$, it is not influenced by the object itself and very similar to the initial pulse, Figure 3.8. The recorded and analytically calculated data are similar, but not match identically. The attenuation at high frequencies is due to the transfer function of the sonar system, which is discussed in Chapter 4 Section 4.1.



(a) experimental data in TD

(b) analytically calculated data in TD

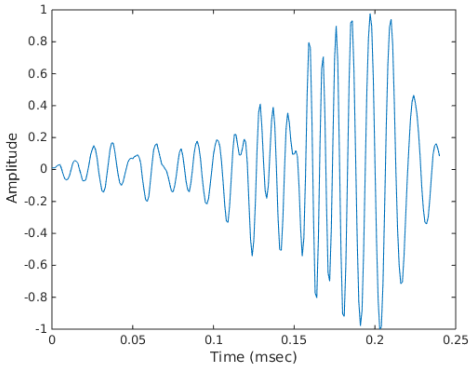
(c) experimental data in FD

(d) analytically calculated data in FD

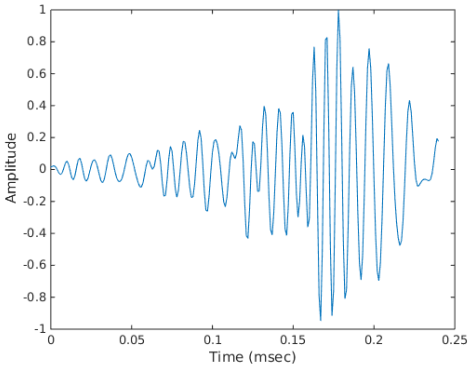
Figure 3.8. Comparison of the analytically calculated and experimental echoes reflected from an aluminium sphere: specular reflection component in time domain (TD) and frequency domain (FD) sampling frequency $1MHz$, FFT size 10001

The secondary reflection, which is following the specular component, is started with the Lamb-wave components, which is a shell-born wave travelling along the shell. The lamb-type component is followed by a number of different components including mid-frequency enhancement, wave travelling through the shell and reflecting from a back wall and others. The components are described in Chapter 5, where some of them are used for the material identification. The biggest secondary reflection in Figure 3.7 appears around $0.3ms$ for both analytical and experimental data. The components in this first set of secondary reflections can hardly be separated for the presented case. The first set of secondary reflection is shown in Figure 3.9. This segment of the echo in time domain looks very alike for experimental and synthetic data, while in frequency domain it can be seen, that the attenuation in high frequencies for the experimental data influences the results. The low frequencies are look alike for both recordings, while in high frequencies some peaks are lost in the experimental

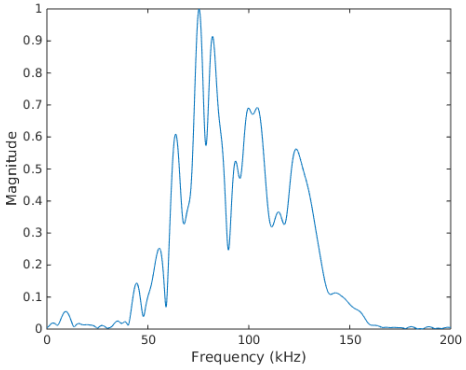
data, which is influenced by the attenuation in high frequencies and difference between ideal sphere and realistic sphere participated in the experiment.



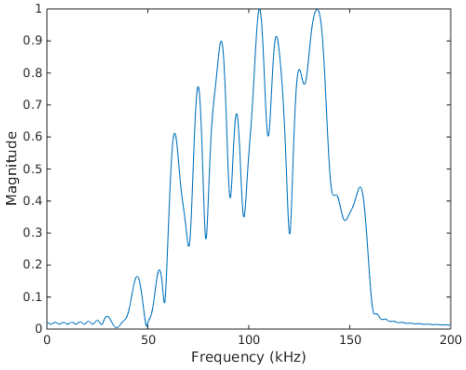
(a) experimental data in TD



(b) analytically calculated data in TD



(c) experimental data in FD



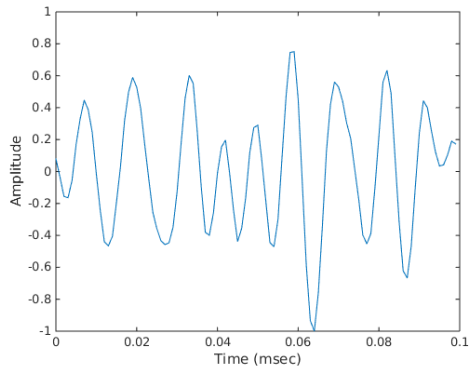
(d) analytically calculated data in FD

Figure 3.9. Comparison of the analytically calculated and experimental echoes reflected from an aluminium sphere: first set of secondary reflections in time domain (TD) and frequency domain (FD) sampling frequency $1MHz$, FFT size 10001

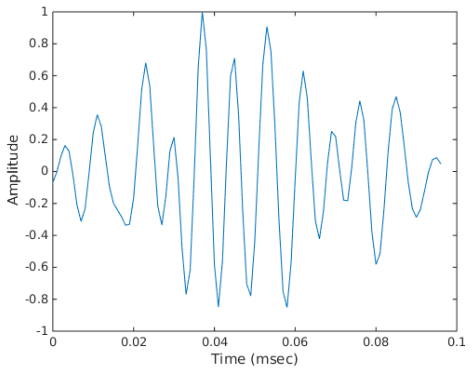
The second separated component of the reflection with a smaller amplitude occurs around $0.5ms$ for both cases, Figure 3.7. This set of the secondary reflections is presented in time and frequency domain in Figure 3.10. It is difficult to compare the components in time domain, while in frequency domain the peaks can be compared. Location of the peaks and notches has the same positioning, but differ in amplitude values. The experimental data is influenced by the attenuation in high frequencies and by abnormalities of the sphere.

In general, based on the presented comparison, it is seen that the pattern of the reflection components and their composition are very alike for the analytically calculated echo and recorded experimental echo. Full description of the components is presented in Chapter 5, where some of these components are used for the shell and filler material identification.

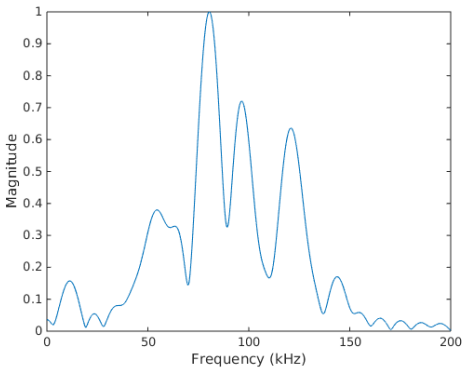
The entire analytical and experimental echoes are presented in frequency domain in Figure 3.11. Both signals are normalised dividing by their maximum value. Focus of the comparison is positioning of the peaks and notches on the frequency range of the initial pulse. Positions of these peaks and notches are related to the object's properties including shape, size, material of the object's shell and material of the object's filler. Therefore these characteristics are most valuable for the target description. The experimental data



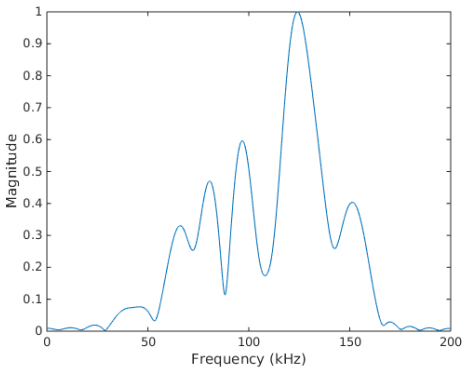
(a) experimental data in TD



(b) analytically calculated data in TD



(c) experimental data in FD



(d) analytically calculated data in FD

Figure 3.10. Comparison of the analytically calculated and experimental echoes reflected from an aluminium sphere: second set of secondary reflectionsin time domain (TD) and frequency domain (FD) sampling frequency $1MHz$, FFT size 10001

has components outside of the frequency range of the initial pulse. These low frequency components are mostly caused by the sonar's hardware (detailed explanation is provided in Section 4.1 Chapter 4) and will not be considered in the comparison. The experimental data has attenuation on the high frequencies. The attenuation is caused by the sonar system itself (Section 4.1) and not taken into account in the comparison. Even though the values of the peaks and notches are not matched identically to each other, but their positions on frequency axis are very alike.

Experimental and numerically calculated echoes are not identical, there are a few reasons for it. In the numerical calculations the sphere is assumed to have a perfect spherical shape, while the real object has three holes, rumpled areas and attached cord for positioning in the water tank. These peculiarities effect the scattering process and as a result the echo itself. In the analytical solution the view angle doesn't influence the result, but has a slight influence into the experimental data since the real spherical object doesn't have a perfect spherical shape. Also the transfer function of the sonar system influence the experimental results. The affect can be observed for high frequencies, especially $130kHz$ and higher. The transfer function of the wideband sonar system is presented in Section 4.1 of Chapter 4.

Despite some difference the positioning of the dips and peaks in frequency domain is similar. Echoes in time domain are also similar to each other, especially from the point

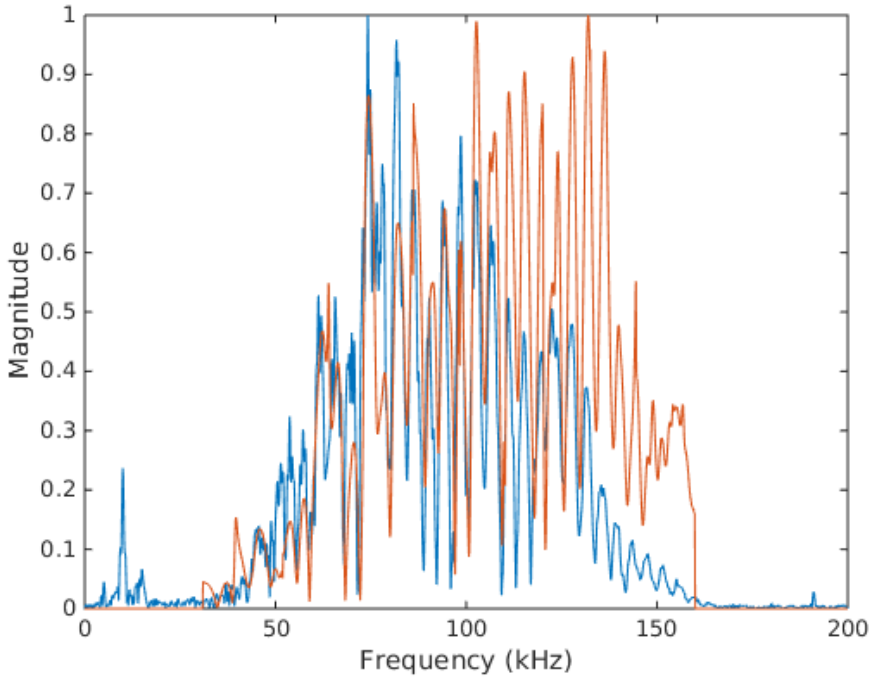


Figure 3.11. Comparison of the entire echo reflected from an aluminium sphere represented in frequency domain: (blue) - experimental data, (red) - analytically calculated data. Recording length in time domain $2ms$, FFT size is 10001, sampling frequency $1MHz$

of view positioning of the specular and secondary echoes. It allows one to conclude that the numerically calculated data simulates the real echo scattering and can be used for data simulation in this work.

3.5 Observation of echoes for spheres with different parameters

It's been already stated that the pulse scattering from an object is changed related to the object's properties. This section presents an observation of the changes in echoes brought by sphere's parameters. The echoes were simulated with only one parameter of the sphere being changed at a time. Therefore influence of this parameter on the scattering can be observed. There are four parameters to be changed: radius, shell thickness, shell material and filler material.

3.5.1 Radius

The target was an aluminium sphere filled with fresh water. The radius of the sphere was changed from $0.05m$ to $0.125m$. Thickness of the sphere equals 2% and it was kept constant for all the radius values. Figure 3.12 presents echoes in time and frequency domains. The signals have sampling frequency $1MHz$, 10001 FFT size and duration $3ms$.

Radius value defines distance which the sound wave travels inside of the sphere and it

has a direct influence into the timing of the reflection components. Timing between specular reflection and secondary reflections increases with radius, as well as between the secondary reflections themselves. It can be shown on the example of the timing between specular reflection and the largest secondary component: for the radius $0.05m$ the timing is $0.2ms$, while for the radius $0.125m$ it is $0.4ms$. Number of secondary components increases with radius: there are no visible components after $1ms$ for the radius $0.05m$, while secondary reflections are still appearing for all $3ms$ length for the radius $0.125m$. Distance between notches in frequency domain is also changed, but in the opposite direction. The distance between the notches decreased as the radius increased: for the radius $0.05m$ the notches spacing is $6kHz$, while for radius $0.125m$ the notches spacing is $2.5kHz$. The radius value is a parameter of the sphere, which has a straight forward influence on the reflected pulse and can be traced in time and frequency domains.

3.5.2 Thickness

For the study of the influence of the shell's thickness, the target was chosen the same - an aluminium sphere filled with fresh water. Radius of the sphere was $0.1m$. Thickness of the sphere was changed from $1\% = 0.001m$ to $4\% = 0.004m$. Figure 3.13 presents echoes in time and frequency domains. The signals have sampling frequency $1MHz$, 10001 FFT size and duration $4ms$.

Thickness of the shell influences the number of secondary echoes, but the timing between them stays constant. Looking at the example of the timing between specular reflection and the largest secondary component - it is $3.5ms$ for all the thickness values. While the thickness increase, secondary components fade slower and amount of secondary components increases: for the thickness 1% the large secondary component are not appear after $1.5ms$, while for the thickness 4% the components can be visible up to $4ms$. In frequency domain the distance between notches increases slightly with thickness values from $3kHz$ for the 1% thicknes to $4kHz$ for the thickness of 4%. Shape of the frequency spectrum is changed along with the thickness, but there is no visible pattern for the changes. The changes brought by the thickness of the shell have visible effect into reflected echo in time and frequency domains.

3.5.3 Shell material

The influence of the shell material was studied for a sphere filled with fresh water. Radius of the sphere was $0.1m$ and thickness was $1\% = 0.001m$. The shell material was changed between four metals, Table 3.2. These materials will be used in material identification approach in Chapter 5 and represent some of the materials available in the Ocean Systems Laboratory, therefore it is useful to study difference between their echoes.

The echoes are presented in Figure 3.15. The signals have sampling frequency $1MHz$, 10001 FFT size and duration $4ms$. Timing of the secondary components doesn't alter much, but the structure of the components differ from material to material. The most significant

Table 3.2. Parameters of the sphere's shell materials

N	material	ρ (kg/m^3)	C_L (m/s)	C_T (m/s)
1	aluminium	2700	6320	3130
2	stainless steel 347	7910	5740	3090
3	copper	8930	4660	2330
4	brass	8560	4280	2300

difference is between aluminium and other metals, due to the significant difference in the parameters of the material: there are more components near to the largest peak of the secondary reflection, the secondary components fade faster, than for other materials. Brass and copper have a very similar parameters and the reflected echoes are look alike in time domain. Frequency components illustrate visible difference between the metals. Positioning of the peaks and their depth is determinative to distinguish between the metals. The difference is more significant for the aluminium, while copper and brass have nearly identical spectrum. Distinguish between shell materials can be challenging if the materials have similar parameters, while materials with distant parameter values can be easily recognised.

3.5.4 Filler material

In this part, the variation due to different filler materials is investigated. The target was an aluminium sphere filled with four different liquids. Radius of the sphere was $0.1m$ and thickness was $1\% = 0.001m$. Parameters of the filler liquids are presented in Table 3.3. These liquids are used in the material identification approach in Chapter 5 as well.

Table 3.3. Parameters of the sphere's filler materials

N	material	ρ (kg/m^3)	C_L (m/s)	C_T (m/s)
1	salty water	1025	1530	-
2	water	1000	1480	-
3	oil	948	1430	-
4	kerosene	810	1320	-

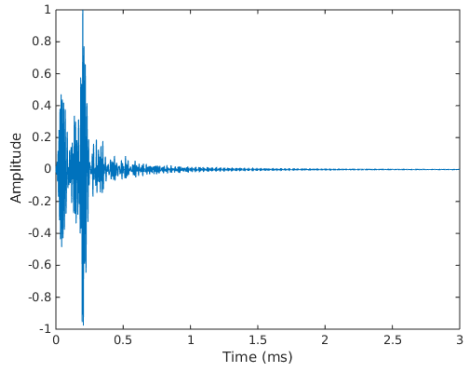
Timing of the secondary components is slightly different. It is increasing from salty water to kerosene: the timing between specular reflection and the largest secondary component increases from $0.32ms$ for salt water to 0.36 for kerosene. Kerosene has a significant difference comparing to other liquids, it has more secondary components, which appear later in time scale, also the largest secondary component has two parts. Significant difference is observed between salty water and kerosene fillers, which can be explained by the largest difference in parameters of the liquids. Difference between frequency spectrum is visible for all the liquids, it is mainly shape of the spectrum, as well position and prominence of peaks and notches. The frequency spectrum of the kerosene filler is bimodal, while water, salt water and oil have unimodal spectrum. Difference in filler materials is visible in time domain with timing of the reflection components and in frequency domains with spectrum shape and position of the peaks and notches.

3.6 Summary

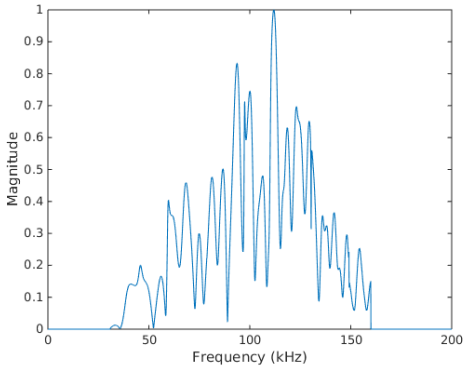
This chapter presented a numerical calculation of a wideband echo reflected from a spherical object immersed in water. The calculations were based on the existed work and adapted for the case of 2-layer sphere with different liquid fillers. Experimental and analytically calculated echoes were compared for an example of an aluminium sphere, radius $0.075m$ and shell thickness $0.001m$, filled with fresh water. Echoes in the time domain are similar to each other, especially in the timing of the primary (specular) and secondary echoes. The primary and two sets of the secondary reflections were compared individually in time and frequency domains. The secondary echoes contain number of components which will be used in Chapter 5 in the material identification approach. It was showed that the experimental and analytically calculated data are similar, but have some differences. The entire echo in frequency domain was studied and despite a small alteration in the positioning of the dips and peaks, the echoes in frequency domain look alike. The differences are due to the imperfection of the spherical shape and influence of the wideband sonar system. Based on the comparison, it has been demonstrated that the numerically calculated data simulates the real echo scattering and can be used for data simulation.

Using the presented approach, echoes reflected from spheres with different physical properties, were generated. This work demonstrated influence of sphere properties into the composition of the reflected pulses. Variation in the echoes due to the difference in the radius, shell thickness, shell material and filler material has been observed for a number of examples. Gain in the radius of the sphere increased the timing between specular reflection and secondary reflections, as well as between the secondary reflections themselves. The distance between notches in frequency domain changed inversely, it decreases while radius increases. The thickness of the shell affects the number of secondary echoes. Also the distance between the notches increased slightly as the thickness was increase. Shell and filler materials influence the structure of the echoes. Frequency components illustrate visible difference between the materials.

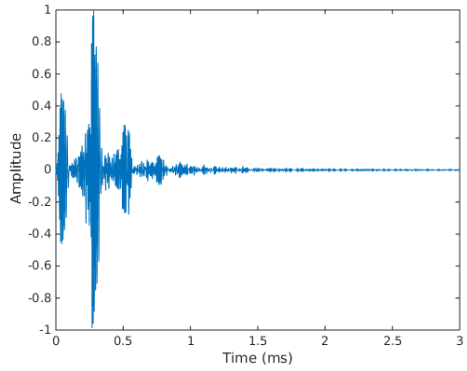
Description of the noticeable changes in the scattering has presented a basis for the techniques developed in Chapter 4 and Chapter 5, where the changes are used for the object classification and material identification.



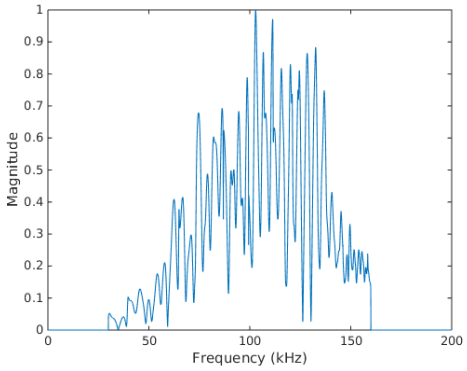
(a) radius 0.05m: TD



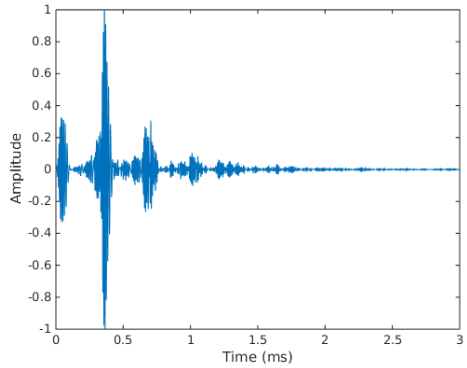
(b) radius 0.05m: FD



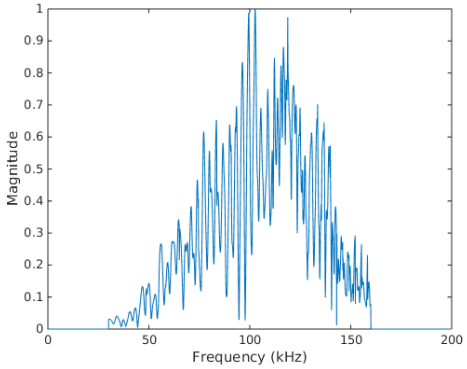
(c) radius 0.075m: TD



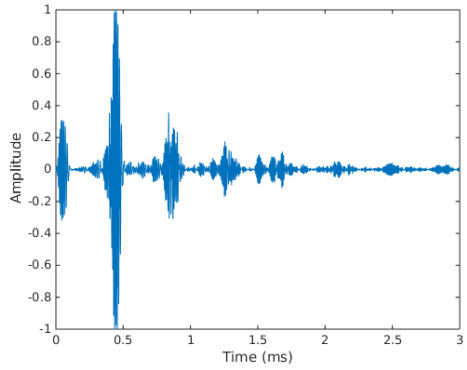
(d) radius 0.075m: FD



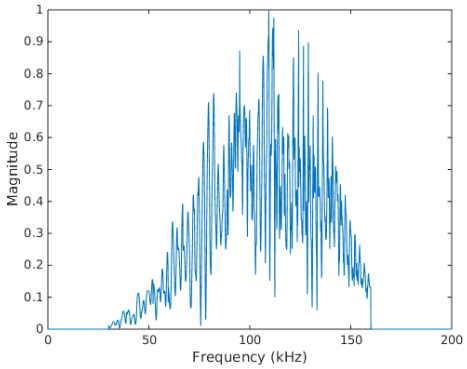
(e) radius 0.1m: TD



(f) radius 0.1m: FD

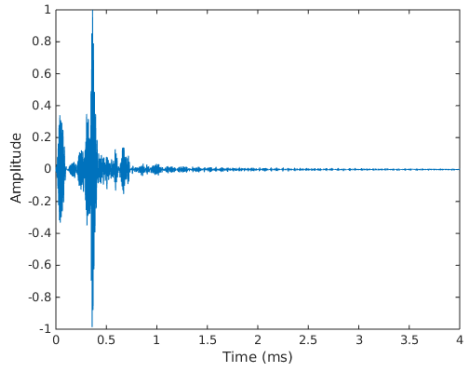


(g) radius 0.125m: TD

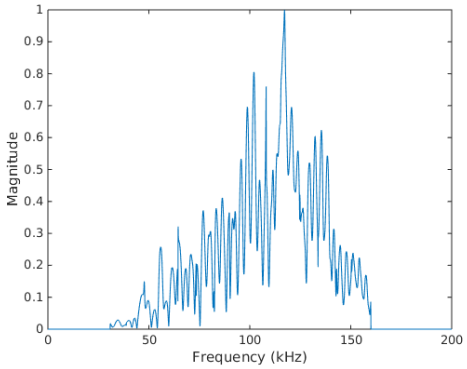


(h) radius 0.125m: FD

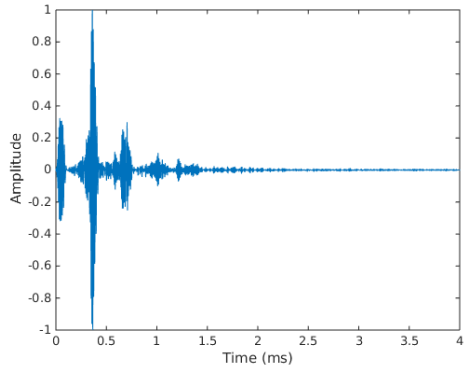
Figure 3.12. Set of synthetic echoes, which illustrates influence of the sphere's radius on the returned echo. The echoes are reflected from aluminium spherical shells with different radii, this is presented in time domain (TD) and frequency domain (FD): duration 3ms, sampling frequency 1MHz, FFT size 10001



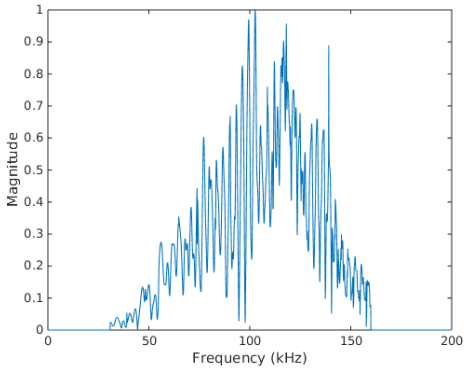
(a) thickness 1%: TD



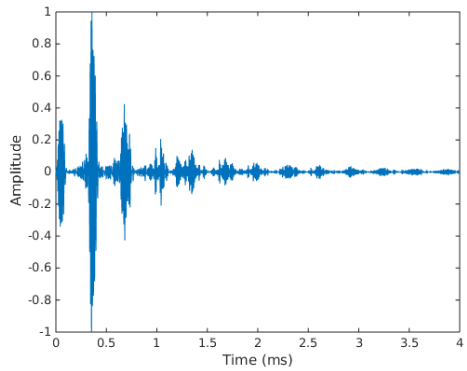
(b) thickness 1%: FD



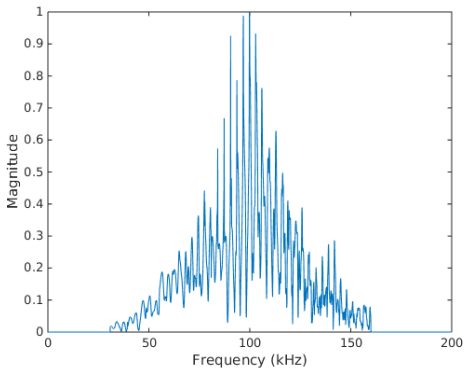
(c) thickness 2%: TD



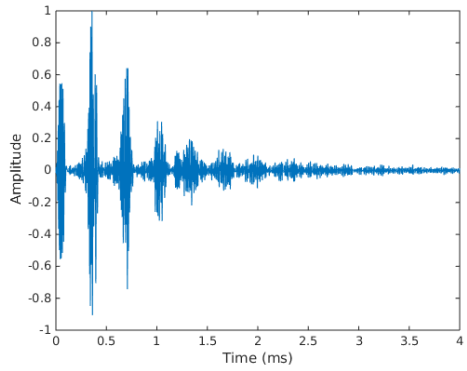
(d) thickness 2%: FD



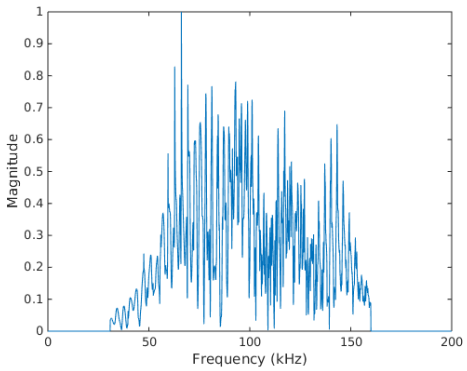
(e) thickness 3%: TD



(f) thickness 3%: FD

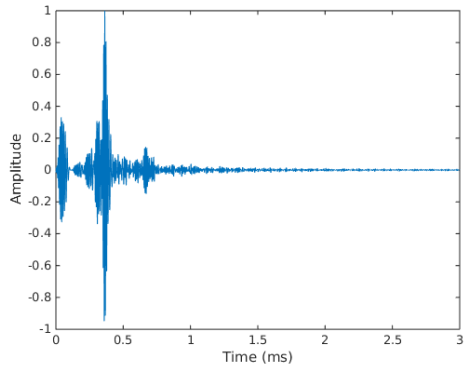


(g) thickness 4%: TD

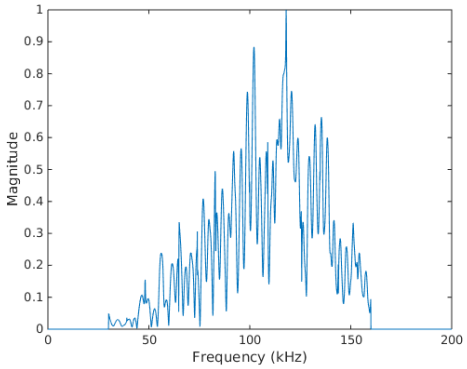


(h) thickness 4%: FD

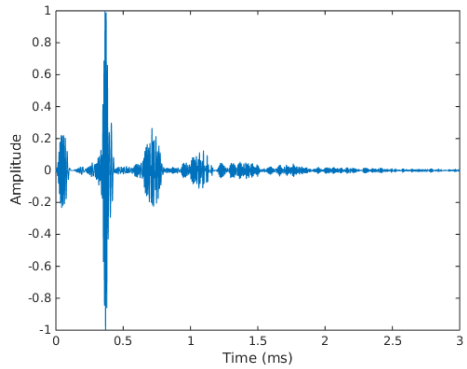
Figure 3.13. Set of synthetic echoes, which illustrate the influence of the shell's thickness on the returned echo. The echoes are reflected from aluminium spherical shells with different shell thickness, this is presented in time domain (TD) and frequency domain (FD): duration $4ms$, sampling frequency $1MHz$, FFT size 10001



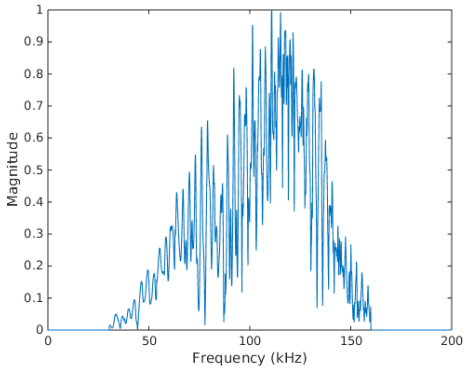
(a) aluminium shell: TD



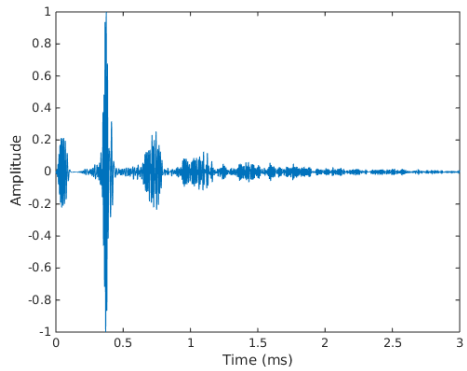
(b) aluminium shell: FD



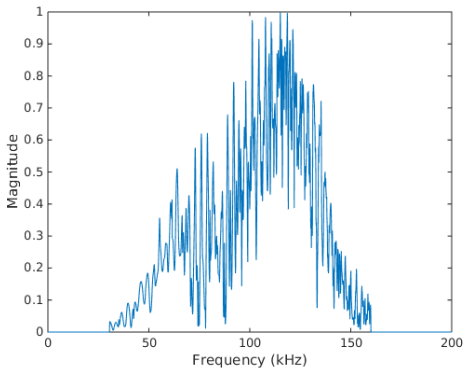
(c) stainless steel shell: TD



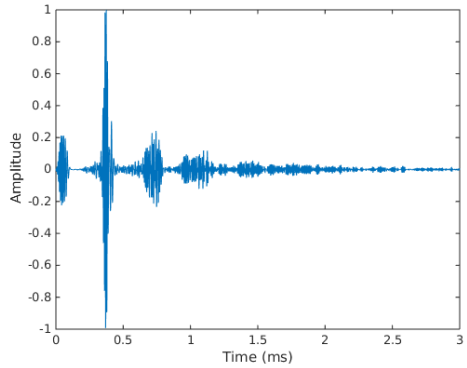
(d) stainless steel shell: FD



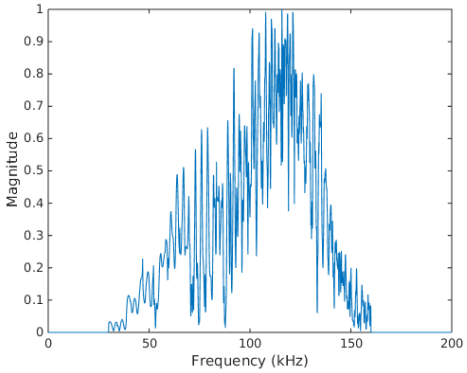
(e) copper shell: TD



(f) copper shell: FD

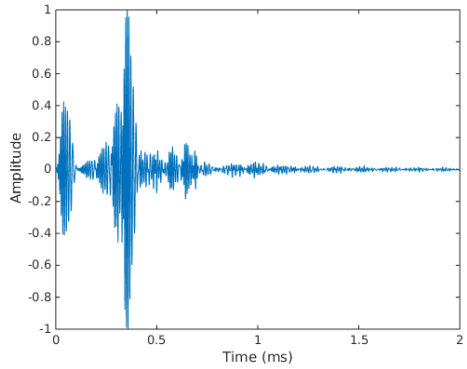


(g) brass shell: TD

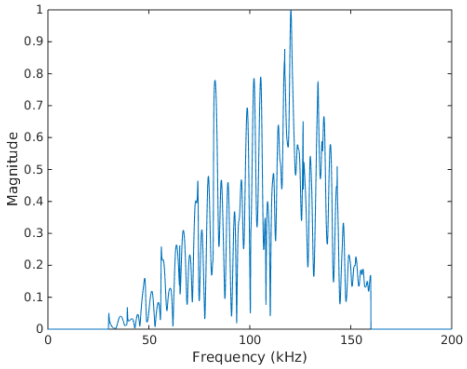


(h) brass shell: FD

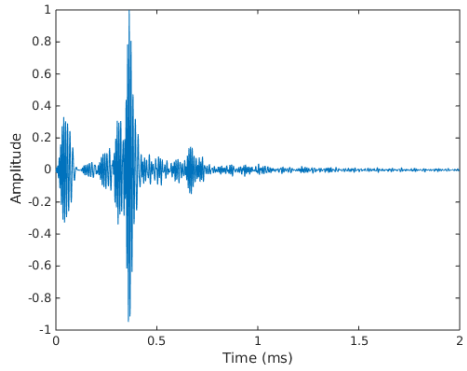
Figure 3.14. Set of synthetic echoes, which illustrates influence of the shell material on the returned echo. The echoes are reflected from spherical shells made of different shell materials, this is presented in time domain (TD) and frequency domain (FD): duration 3ms, sampling frequency 1MHz, FFT size 10001



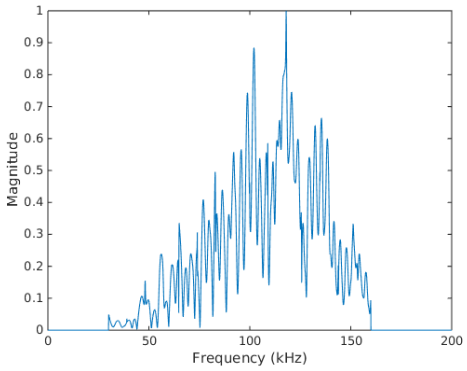
(a)salt water: TD



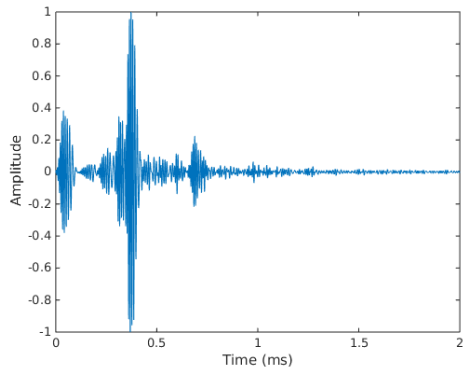
(b)salt water: FD



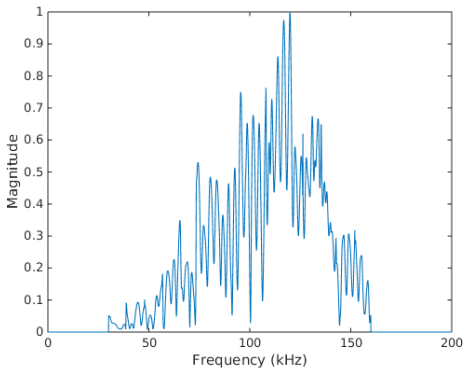
(c) water: TD



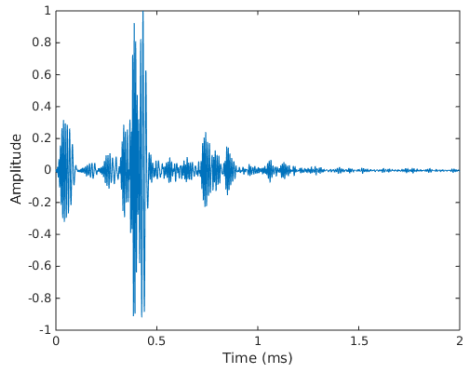
(d) water: FD



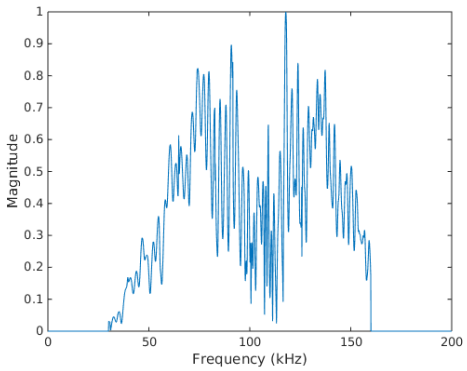
(e) oil: TD



(f) oil: FD



(g) kerosene: TD



(h) kerosene: FD

Figure 3.15. Set of synthetic echoes, which illustrates influence of the filler material on the returned echo. The echoes are reflected from aluminium spherical shells filled with different filler materials, this is presented in time domain (TD) and frequency domain (FD): duration $2ms$, sampling frequency $1MHz$, FFT size 10001

Chapter 4

Classification using machine learning techniques

A learning experience is one of those things that say, "You know that thing you just did? Don't do that".

Douglas Adams, The Salmon of Doubt

This chapter provides solution for object classification using wideband echoes. The objects are spherical shells filled with liquids or an air. Material of the shell and filler, thickness of the walls are considered as physical properties of the sphere. In this way, all the objects have the same spherical shape and differ in physical properties and size. In Chapter 3 Section 3.5, the variations in the reflected pulses, related to the sphere's size, thickness, filler and shell materials, were presented. These distinctions are used in this chapter for the object classification using machine learning techniques. Each sphere has its unique combination of the shell thickness, filler and shell materials, radius, and hence it has a unique reflected echo, which represents the target.

The classification results are presented for the experimental data. The description of the experimental set-up is given in Section 4.1. Firstly, in Section 4.2, the objects are classified with a multi-class classifier, where one class represents one object [98]. Spherical objects with different physical properties are classified based on the time-frequency representation (TFR) of their echoes. The next step towards the recognition of the physical property is classification of objects based on their filler [99], which is presented in Section 4.3. The objects are classified into two classes, where one class is water filled spheres and another class is air filled spheres. The objects are described by a form function [97], which expresses a pressure field scattered from a target in a range of frequencies. In this way, the machine learning techniques are applied for the material based classification.

The wideband signal and machine learning techniques can be used not only for an object classification, but also applied to a wider range of tasks, for instance, sea floor recognition. As an example, Section 4.4 presents sea floor recognition using wideband pulses [100]. A citadel constructions and sediment regions in Tallinn bay, Estonia are classified based on

the frequency components of the reflected pulse.

4.1 Experimental environment

The experimental data is recorded in the test tank of the Ocean Systems Laboratory in Heriot-Watt University. The test tank is a closed controlled environment. In open water, there would be range dependant losses and background noise which consist of ambient noise and other interfering noises. They would have an impact into the received pulse and therefore influence results for the approaches in Chapter 4 and Chapter 5. This section describes experimental set-up of the sonar system in the test tank and presents echo selection from the recordings, which is required for the object classification.

4.1.1 Experimental set-up

The echoes are recorded in a $3m \times 4m \times 2m$ water tank, Figure 4.1. The objects are hung in the water using weight and placed in front of the sonar's transducer. The sonar is a wideband sonar produced by Hydrason. The length of the tank is $4m$ and the object is placed in $(1-3)m$ away from the sonar's transducer (distance r).

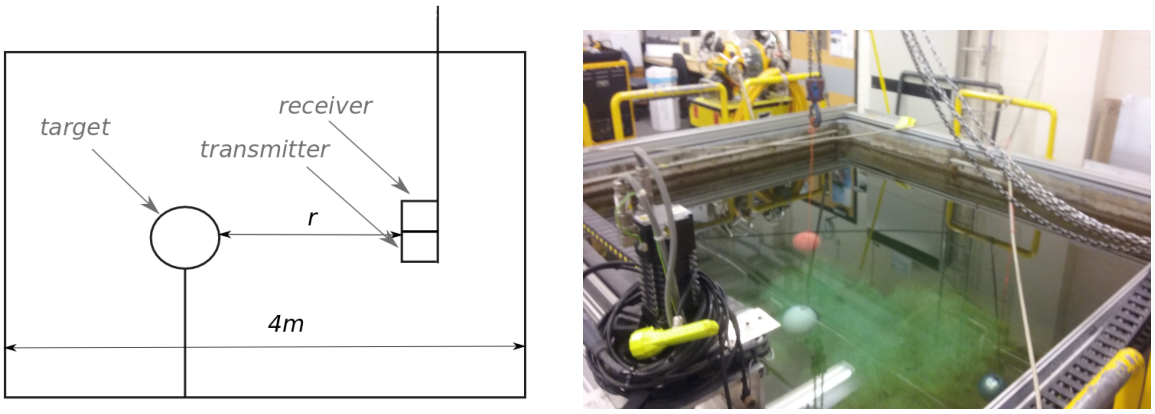


Figure 4.1. Experimental set-up in the test tank

The wideband sonar operates in the frequency range from $20kHz$ to $180kHz$ with sampling frequency $1MHz$. The range of the sonar is $75m$ and the beam width of the receiver is presented in Table 4.1 [101].

Table 4.1. Beam width of the receiver defined by the -6dB point

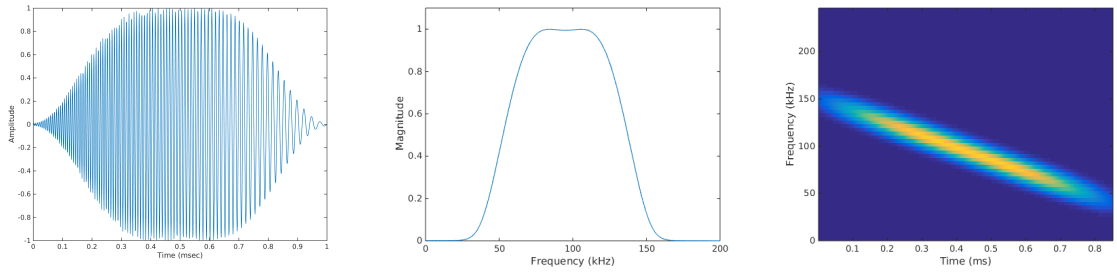
Frequency	40	40	80	80	120	120
Receiver orientation	H	V	H	V	H	V
Beam width	48°	96°	24°	48°	16°	32°

The range of the sonar's far field can be calculated by Equation 4.1 [102].

$$R \geq \frac{2D^2}{\lambda}, \tag{4.1}$$

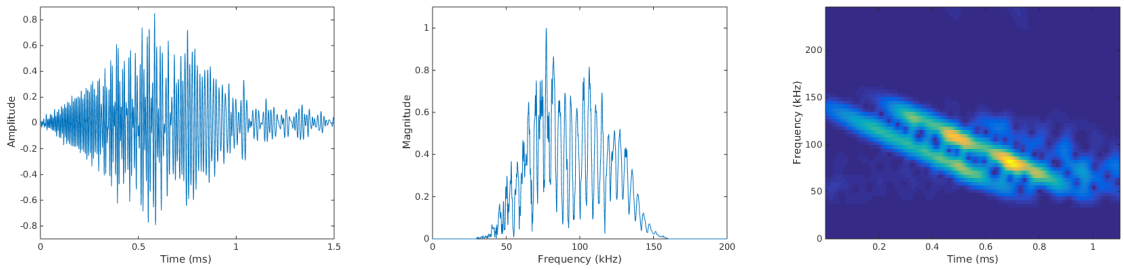
where R is the range, D is the diameter or length of the transducer, λ - the wavelength of the transmitted pulse. The size of the sonar's transducer is $2.51cm \times 4.82cm$. The length of the transducer is used for the calculation of the far field range, in this way $D = 4.82cm$. The wavelength of the sonar's frequency range $(20-180)kHz$ in fresh water is $(0.0740-0.0082)m$. Using Equation 4.1, the far field range for the wideband sonar is $R \geq 0.61m$. In this way, all the objects have to be in a distance $r > 0.61m$ to be in the far field of the sonar.

The wideband sonar allows transmitting pulses of different shape and duration. Linear down chirp pulses were used in this study. Figure 4.2 shows an example of the initial chirp pulse duration $1ms$, which was fed to the transducer. The pulse is generated in the frequency range $(160-30)kHz$ and the 3 dB bandwidth of the pulse is $(136-52)kHz$ due to the Blackman windowing of the signal [103]. The pulse which is sent to the target differs from the initial generated signal. The difference between the generated signal and the pulse which reaches the target is described in Section 4.1.2.



(a) TD (c) FD (c) TFD

Figure 4.2. Initially generated pulse fed to the transducer - linear chirp, duration $1ms$. The sampling frequency of the signal $1MHz$, FFT size 10001, the signal is presented in TFD using STFT with Gaussian windowing and FFT size 1000



(a) TD (c) FD (c) TFD

Figure 4.3. Echo from an aluminium sphere filled with fresh water. The sampling frequency of the signal $1MHz$, FFT size 10001, the signal is presented in TFD using STFT with Gaussian windowing and FFT size 1000

The reflected signal from the target is changed. Figure 4.3 presents a pulse, reflected from an aluminium sphere filled with fresh water. Chapter 3 introduced different components of the reflected pulse. The duration of the initial pulse ($0.1ms$) doesn't allow to separate reflection components in time domain. In time-frequency domain, two separate chirps can be identified. The first chirp is related to the specular reflection from the sphere,

while the second chirp is the reflection from the back wall of the sphere. All the other components of the reflection, including Lamb-type wave, mid frequency enhancement and others are overlapped and cannot be easily distinguished. The changes in the pulse are due mostly to the interaction with the object, but the response and transfer function of the sonar system characteristic also influences the returned pulse.

4.1.2 Wideband sonar

The influence of the sonar system can be studied through the system transfer function. It allows one to separate the changes brought by the equipment from changes due to the reflection from an object.

The sonar system can be represented by a scheme in Figure 4.4. The initial generated digital pulse $u[n]$ goes into the digital-to-analogue converter (DAC). The analogue signal is augmented by the power amplifier (PA) and transmitted to the medium by the transmitter T_x. In the medium ("Acoustic system" block), the pulse propagates, reflects from a target and returns to the receiver, R_x. The acquired signal is amplified by the pre- amplifier (*Pre-A*) and converted into its digital form by the analogue-to-digital converter (ADC) [104].

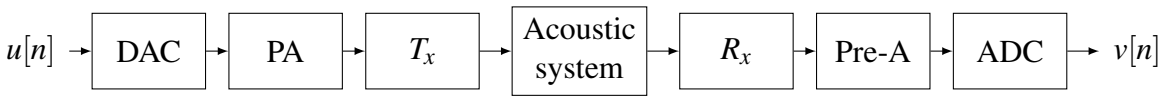


Figure 4.4. Scheme of the sonar system

The system has linear and non-linear effects on the return pulse. Each component of the sonar system influences the signal. This influence can be described by a transfer function. In this way, the output signal from the DAC can be described, as $u_1(t) = T_{DAC}\{u[n]\}$, where T_{DAC} is a transfer function of the DAC, and output of PA is $u_2(t) = T_{PA}\{u_1(t)\}$ with T_{PA} is a transfer function of the PA, etc. The output signal of the sonar system is influenced by each block of the scheme. The influence of the sonar system can be describe by Equation 4.2, [105].

$$v[n] = T_{ADC}\{T_{Pre-A}\{T_{R_x}\{T_{T_x}\{T_{PA}\{T_{DAC}\{u[n]\}\}\}\}\}\}, \tag{4.2}$$

where $u[n]$ is an initial digital pulse, $v[n]$ - returned pulse after analogue -to-digital transformation and T is a transfer function of each block of the sonar system. This equation presents the impact of the sonar system on the output signal $v[n]$ through the transfer functions of the system's components. The system transfer function is a merge of the components transfer functions. This influence of the system can be observed by an experiment where the flat motionless water surface is insonified with a transmitted signal. Due to the high impedance boundary, the water surface is a pressure realise surface. This results in a pure reflection with a 180° phase shift. In this case, all the changes in the pulse are due to the transfer function of the sonar system. The experimental set-up is similar to the described above, aside from the transducer orientation. The transducer is placed looking towards the

water surface, Figure 4.5.

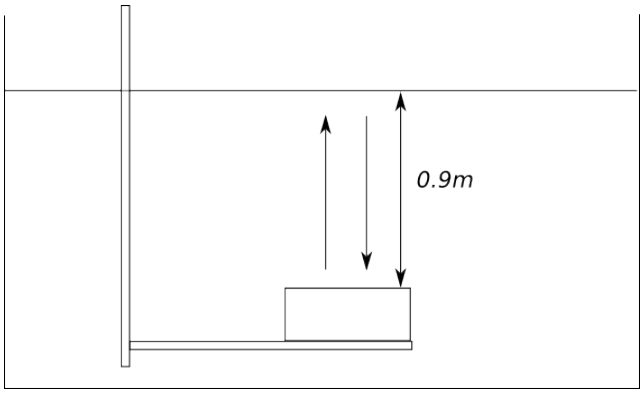


Figure 4.5. Transducer orientation for the experimental set-up

Initial (transmitted) and reflected pulses are presented in Figure 4.6. The phase of the reflected pulse is inverted due to the reflection from the water surface. In the frequency domain, it is observed, that the high frequency region is influenced the most by the system with a visible attenuation from $85kHz$. This attenuation is expected to be in all the recordings. In low frequency region of the reflected pulse ($f < 25kHz$), there are components, which are related to the sonar system noise. These components are also expected in all the recordings.

4.1.3 Echo selection

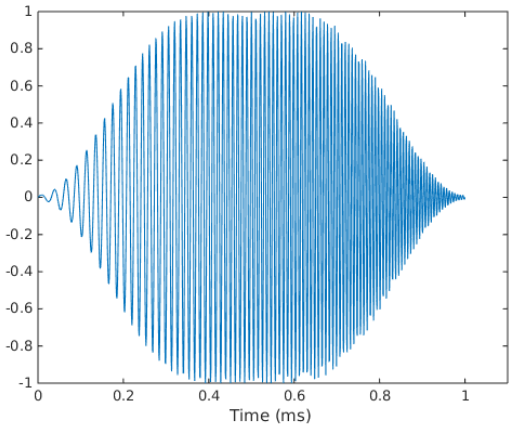
The original recorded response in the test tank contains number of reflections, including reflection from the target, walls, bottom of the tank and other surfaces, Figure 4.7.

In the presented case, the first component is a recording of the transmitted pulse, the second component is the reflection from the target, the third one - reflection from the back wall of the test tank, etc. The reflection from the target contains the only useful information for classifiers. This segment is selected from the recording based on the matched filtering of the initial and returned pulses. The surfaces, the pulse reflected from, are presented as peaks on the filter's output, Figure 4.8. The figure presents first $0.03s$ of the recording.

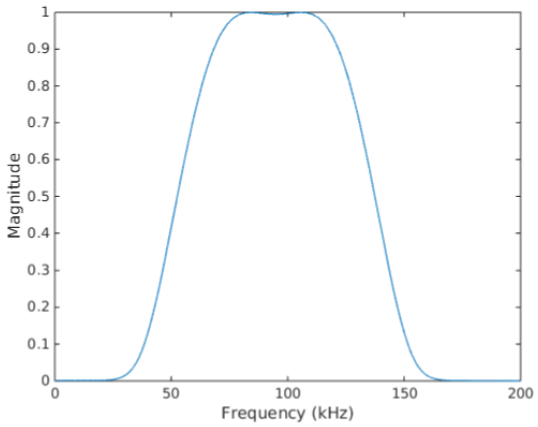
The peak related to the object of interest is located in the known range between $1m$ and $3m$, corresponding to time of $1.3ms$ and $4ms$ propagation in fresh water. The first significant peak in the range is chosen as the target's surface. The starting point of the scattering, t_{sc}, is calculated based on the peak position, t_{peak}, and duration of the initial pulse, Δt_{pulse} , Equation 4.3.

$$t_{sc} = t_{peak} - \Delta t_{pulse}/2 \tag{4.3}$$

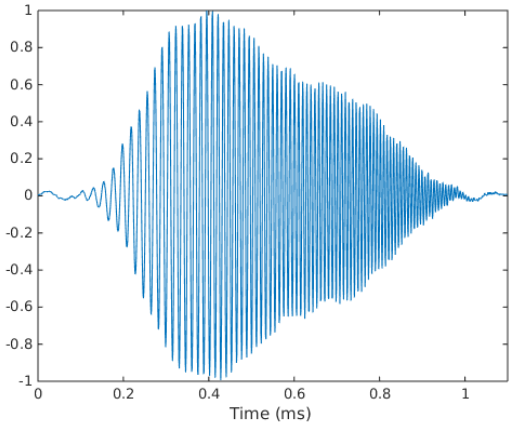
The shift of the position on half of the initial pulse length ($\Delta t_{pulse}/2$) is due to the convolution of the return echo with the time-reversed initial pulse [109]. The originally generated pulse is symmetrical, while the pulse, which reach the target is not, due to the influence of the sonar system (Section 4.1.2). The asymmetry of the pulse doesn't influence



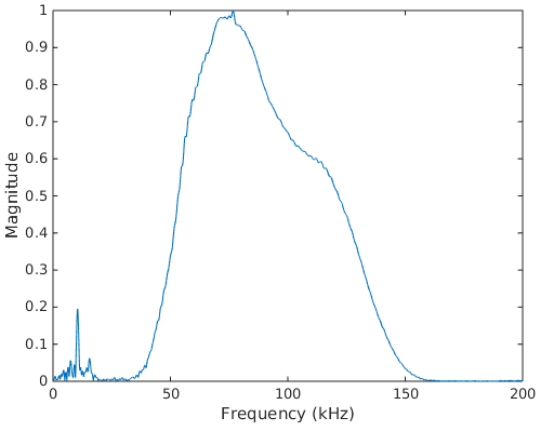
generated pulse in TD



generated pulse in FD



reflected pulse in TD



reflected pulse in FD

Figure 4.6. Normalised generated pulse and reflected echo from the water surface with sonar system noise at low frequencies and attenuation at high frequencies. Pulse duration $1ms$, sampling frequency $1MHz$ and 10001 FFT step

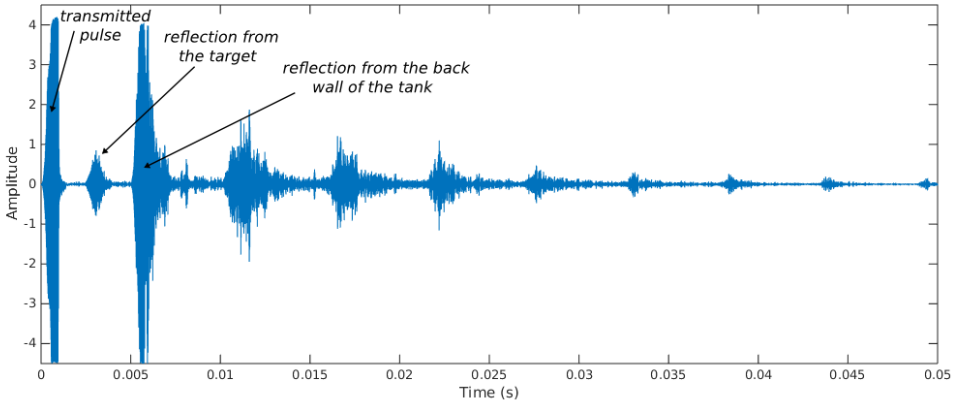


Figure 4.7. Recording of a response: sampling frequency $1MHz$, recording length $0.05s$

the positioning of the peak on the matched filter output. The selected reflection from the target is presented in Figure 4.9, where the the starting point of the scattering is calculated using Equation 4.3. The calculated starting point locates the reflected echo at the beginning of the segment, therefore Equation 4.3 can be used for the echo selection. The duration of

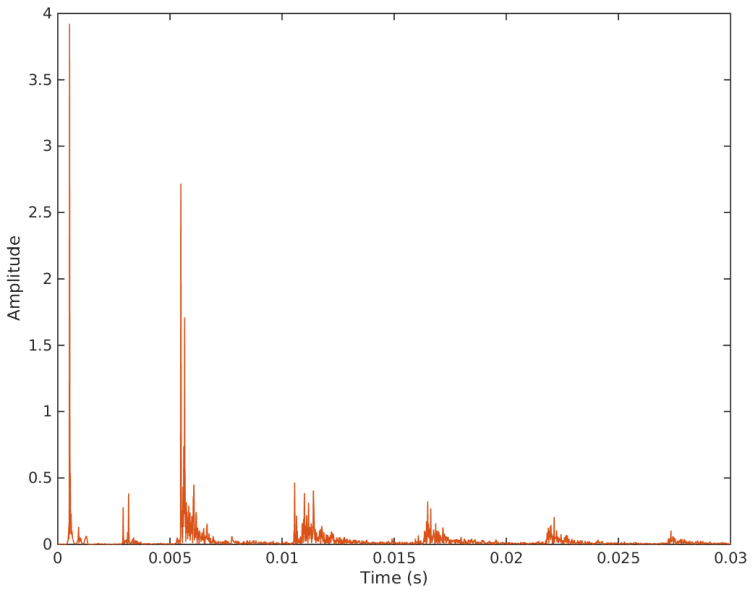


Figure 4.8. Matched filter's response. Recording has $1MHz$ sampling frequency and duration $0.03s$

the scattering sample is fixed for all experiments and equals $2ms$. The value is chosen based on the duration of the initial pulses and geometry of the test tank.

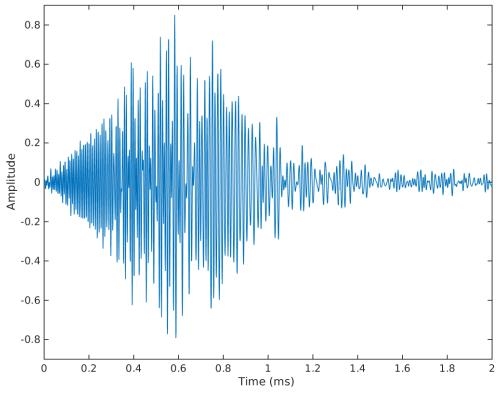


Figure 4.9. Selected echo reflected from the target: duration $2ms$, sampling frequency $1MHz$

The echo selection is essential for the classification, as only the reflection from the target has to be in the selected segment. Otherwise the classifier would be fed with a misleading data and classification will be compromised.

4.2 Classification of spherical objects with different physical properties

In this section, spherical objects with different physical properties are classified based on the time-frequency representation (TFR) of their scattering pulses. There are a number of publication, which use wideband pulses for object classification. For example, Gaunaurd

et al. [4] insonified a sphere with recorded wideband dolphin pulses and define specific features of the echoes connected to the physical properties of the sphere. Pailhas *et al.* [32] classify seven different objects based on the location of the notches in frequency representation of their wideband echoes. Qiao *et al.* [33] apply singular value decomposition for echoes in time-frequency domain (TFD) and use support vector machine to classify three copper cylindrical shells with different thickness. In contrast to the previous approaches, machine learning techniques are applied to classify objects with the same shape and different physical properties. The classification can be described by a three step process, Figure 4.10.

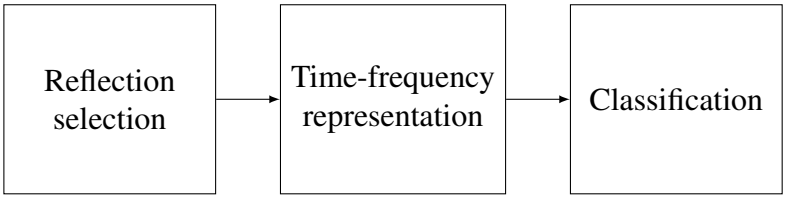


Figure 4.10. Scheme of the classification approach

The reflection selection is described in Section 4.1. The selected pulse is represented in time-frequency domain (TFD) using short time Fourier transform (STFT). Images of the STFT magnitude are fed to a number of classifiers: convolution neural network (CNN), multilayer perceptron (MLP), support vector machine (SVM) and gradient boosting (GB). Different window length of STFT are applied to investigate the influence of the time and frequency resolution into the performance of the classifiers. This section presents multi-class classification where each class presents one target with it's own unique characteristics.

4.2.1 Echo representation

The time-frequency representation (TFR) of a signal provides the distribution of its energy over time and frequency simultaneously [71]. It describes how a signal's frequency composition changes with time. There are many different techniques which transform a signal from time domain into the time-frequency domain (TFD) [64]. In this work, the magnitude of the short time Fourier transform (STFT) is used for the TFR. The signal in time domain is segmented into a sequence of short sections using a window function $w(n)$. Each segment is transformed into the frequency domain using the Fourier transform.

The magnitude of the STFT can be displayed as 2D plot with frequency and time axis, Figure 4.11. The TFD images are presented in colour for visualisation purposes, while for classification, a single channel (grayscale) images were used. The size of the TFD image fed to the classifier is fixed to 50×75 pixels. The value of each point on the graph represents an energy of a particular time interval and frequency. There are a few reasons why the short time Fourier transform was chosen for the TFR. Firstly, STFT doesn't contain a cross-product, which is common for the Wigner-Ville distribution. Secondly, the approach is a trade-off between time and frequency resolution, which allows changing the resolution to

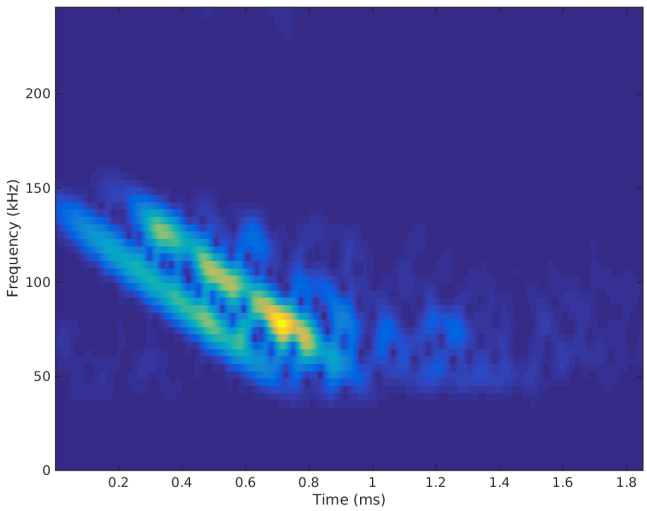


Figure 4.11. Echo from an aluminium sphere filled with water presented by a magnitude of the STFT: recording length $2ms$, sampling frequency $1MHz$, 2000 FFT size and Gaussian windowing, window length $0.15ms$

investigate an influence of the time and frequency resolution into the performance of the classifier.

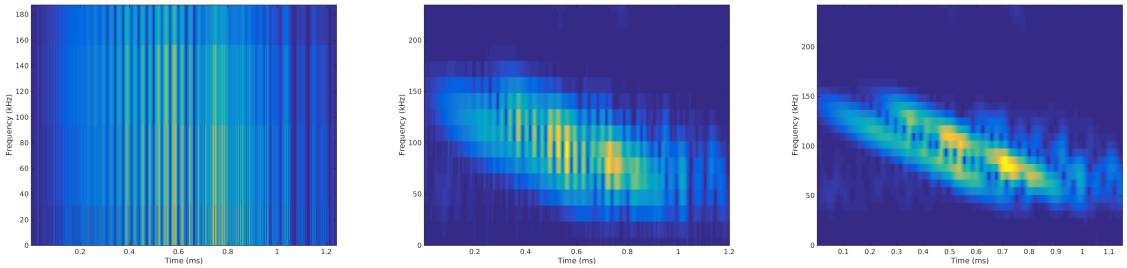
The resolution in time is evaluated by the length of the window, L. A segment with a length L is presented by a single vector on frequency axis. Decreasing the window length L leads to improvement in time resolution, Δt. The frequency resolution depends on the sampling frequency f_s and number of samples in the window N, Equation 4.4.

$$\Delta f = \frac{f_s}{N} \tag{4.4}$$

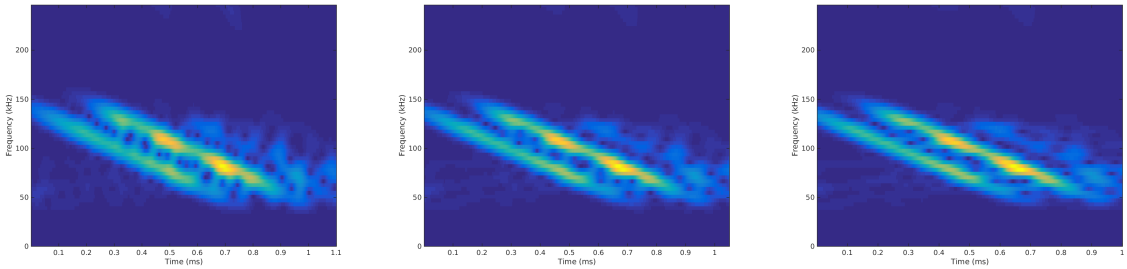
Increasing the window length improves the frequency resolution. In this way, the window length is a compromise between resolutions of time and frequency axis. The length of the scattering segment equals $2ms$ for all the experimental data with sampling frequency $1MHz$, 2000 FFT size and Gaussian windowing. Figure 4.12 illustrates influence of the STFT window length on the time and frequency resolution. The signal is an echo from an aluminium object. The window length, L, is varied from $0.01ms$ to $0.4ms$, while the step size, R, is kept constant. It gives a variation in frequency resolution from $100kHz$ to $2.5kHz$. Based on the visual observation of the TFR images, the major window length for the classifier was chosen equals $0.15ms$, it presents a good compromise for the time and frequency resolution. The architecture of the neural network classifiers are built for the example of $0.15ms$ window length and then trained and tested for all the other values as well.

4.2.2 Classification classes

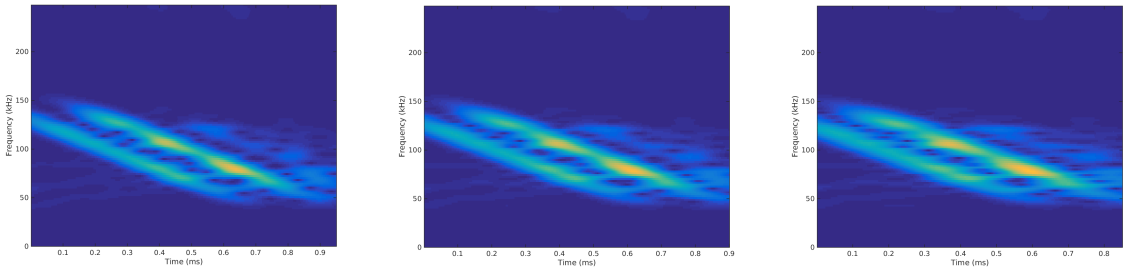
The pulses represented in TFD are classified into five classes with four objects and one empty scene, Table 5.9. One class is represented by one object. The objects are presented



(a) L = 0.01ms (b) L = 0.05ms (c) L = 0.10ms



(d) L = 0.15ms (e) L = 0.20ms (f) L = 0.25ms



(g) L = 0.30ms (h) L = 0.35ms (i) L = 0.40ms

Figure 4.12. Magnitude of the STFT with different window length L: recording length $2ms$, sampling frequency $1MHz$, 2000 FFT size and Gaussian windowing

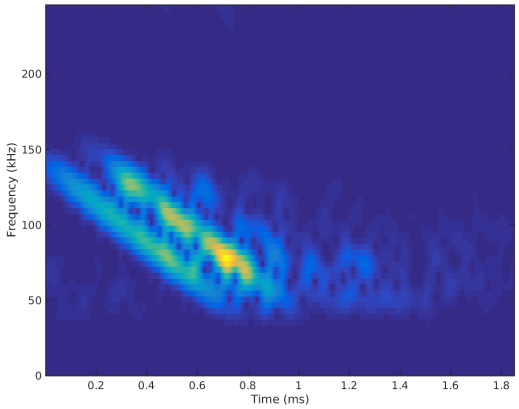
by one aluminium sphere and three plastic spheres. The empty scene doesn't contain any objects within the expected range from $1m$ to $3m$. The objects have the same spherical

Table 4.2. Description of the objects

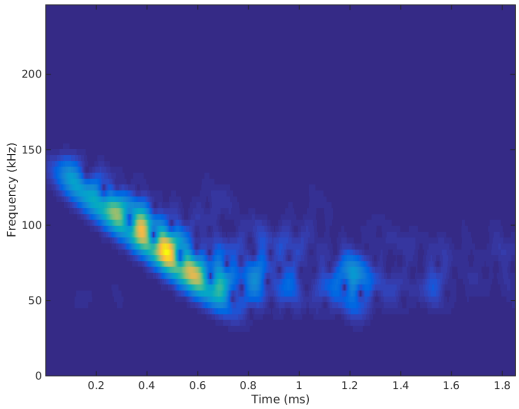
classes	object	diameter, m	shell material	filler
class 1	sphere	0.15	aluminium	water
class 2	sphere	0.274	plastic	air
class 3	sphere	0.208	plastic	water
class 4	sphere	0.208	plastic	air
class 5	no objects	-	-	-

shapes and differ by radius, thickness of the shell, filler and shell materials. The parameters of a spherical target influence the scattering and create specific changes in the initial pulse. The spherical shape of the objects eliminates the influence of the view point on the scattering. Each object is insonified with linear chirp pulses. The duration of the pulses changes from $0.1ms$ to $2ms$. Figure 4.13 illustrates scattering of $1ms$ pulse in TFD for each class. The length of the scattering segment is $2ms$ with sampling frequency $1MHz$, 2000 FFT

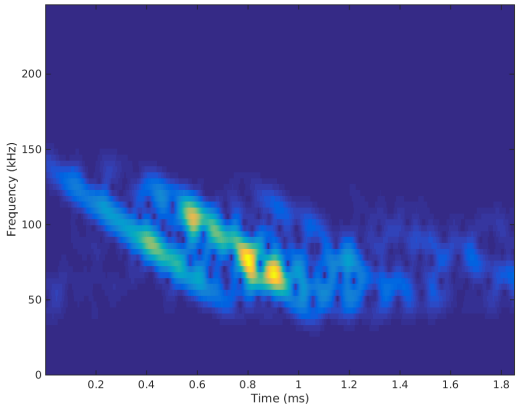
size and Gaussian windowing. STFT is implemented with window length 0.15ms, which is chosen as major for the classifier.



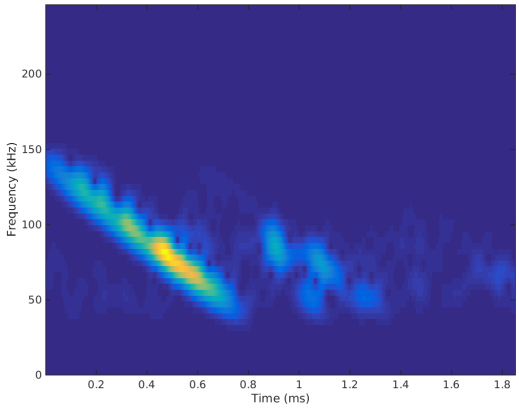
class 1



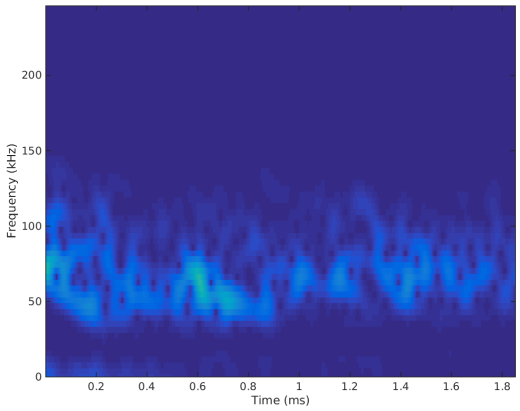
class 2



class 3



class 4



class 5

Figure 4.13. TFR of the reflected echo for each class: class 1 - aluminium sphere diameter 0.15m filled with water; class 2 - plastic sphere diameter 0.274m filled with air; class 3 - plastic sphere diameter 0.208m filled with water; class 4 - plastic sphere diameter 0.208m filled with air; class 5 - empty scene. Recording length 2ms, sampling frequency 1MHz, 2000 FFT size, Gaussian windowing, window length 0.15ms

4.2.3 Classification

For the TFR based object classification four different classifiers were applied and compared: convolution neural network (CNN), multilayer perceptron (MLP), support vector machine (SVM) and gradient boosting (GB). The use of different neural network classifiers, namely a convolutional neural network, and a multilayer perceptron (MLP) was motivated by a good results in the image-based classification for the neural networks. SVM and GB are common used techniques with a substantial body of theory behind them.

In order to provide a fair evaluation, a 5-fold cross-validation was performed, and at each fold, 5 instances of the same neural network were trained, with different random initializations of their weights. This includes the effect of both data and random initialization on classification accuracy. Architectures of the neural networks and parameters of the classifier are described below for each classifier separately.

Convolution neural network

This convolution neural network is constructed from two 2D convolutional layers with following max-pooling layers after each convolutional. The 2D convolutional layer contains n square filters of size s (Conv2D(n, s)). The max-pooling layers have subsampling size s (Max-Pool(s)). Two fully connected layers with n output neurons (FC(n)) are used at the end of the network.The network architecture is shown in Figure 4.14.

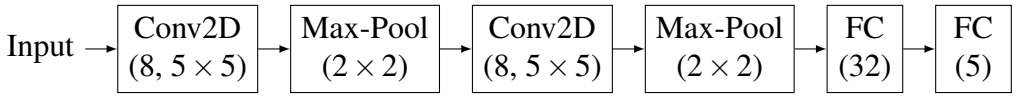


Figure 4.14. Convolutional neural network architecture

All layers use ReLU activation (Equation 4.5), except the last layer that uses a softmax function (Equation 4.6), [93].

$$f(x) = \max(0, x) \tag{4.5}$$

$$f(x) = \left[\frac{e^{x_i}}{\sum_j e^{x_j}} \right]_i \tag{4.6}$$

This network takes a 50x75 input image of the STFT magnitude. It has 36973 parameters for the one-channel grayscale image.

Multilayer perceptron

The multilayer perceptron is contracted from three fully connected layers with configuration: FC(64)-FC(64)-FC(5). The input to this network is a flattened image vector of size $c \times 50 \times 75$, where c is the number of channels in the input image. The model architecture is shown in Figure 4.15.

To prevent overfitting, FC layers are followed by dropout layers [106] with $p = 0.5$, except for the output layer. ReLU activations (Equation 4.5) are used through the network,

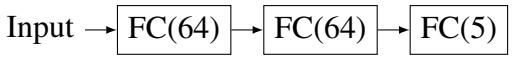


Figure 4.15. Multilayer perceptron architecture

except for output layers that use the softmax activation instead (Equation 4.6).

Both networks are trained with stochastic gradient descent (SGD) for 15 epochs with a learning rate $\alpha = 0.01$ and a batch size $B = 128$. The loss function that is minimized during training is the Categorical Cross-Entropy:

$$L(y,\hat{y}) = -\sum_{i=0}^{N}\sum_{c=0}^{C} y_i^c \log \hat{y}_i^c, \tag{4.7}$$

where N is the number of elements in the dataset, and C is the number of classes ($C = 5$ in this work).

Support vector machine

Multi class support vector machine (SVM) with linear kernel and regularisation coefficient $C = 10$ is applied for the classification. The regularisation coefficient determines a compromise between misclassification and hyperplane complexity [86].

Gradient boosting

Gradient boosting (GB) classifier is chosen with 100 stages, learning rate $\alpha = 0.1$, and maximum tree depth of seven. Both classifier configurations were obtained by grid search on a validation part of the dataset.

The described solution is implemented in Python.

4.2.4 Results

The window length was selected in the range $L \in [0.01, 0.05, 0.10, 0.20, 0.25, 0.30, 0.35, 0.40]$. For each value of L the mean and standard deviation of accuracy over the 25 trained networks (5 folds times 5 network instances) are reported. The results for all the classifiers are summarized in Table 4.3. Bold text in the table is used to highlight largest accuracy for each window length. There are raws where results for two classifiers are highlighted, which means that maximum mean accuracy of one classifier is less than a mean accuracy plus standard deviation of another classifier, therefore both results are highlighted.

Two comparisons are performed. The first is the use of a CNN, MLP, SVM, and GB classifiers. The second is the variation of the window length parameter L. The classifiers are trained using the same cross-validation methodology and report mean and standard deviation of test accuracy. The results are also presented in Figure 4.16, it visualises the mean accuracy and standard deviation for each classifier and provides more interpretable results representation.

Table 4.3. Performance of the CNN, MLP, SVM and GB classifiers for a range of window lengths L

L, ms	0.01	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40
Δf. kHz	100	20	10	6.7	5	4	3.3	2.8	2.5
Mean Accuracy ± Standard Deviation, %									
CNN	**97.4 ± 1.8**	**96.2 ± 1.5**	**98.4 ± 0.8**	**97.9 ± 2.8**	93.3 ± 13.2	**90.7 ± 19.9**	47.2 ± 34.2	**97.4 ± 1.0**	**96.8 ± 2.3**
MLP	60.3 ± 18.1	55.8 ± 15.4	51.3 ± 17.8	57.2 ± 17.7	46.4 ± 17.1	46.9 ± 16.0	51.4 ± 21.5	55.6 ± 15.5	59.1 ± 14.4
SVM	96.9 ± 0.04	95.8 ± 0.14	98.4 ± 0.1	**98.5 ± 0.02**	**98.9 ± 0.03**	**98.0 ± 0.12**	**98.8 ± 0.14**	96.94 ± 0.2	96.0 ± 0.12
GB	94.9 ± 0.08	94.0 ± 0.24	97.3 ± 0.2	97.3 ± 0.23	97.2 ± 0.14	96.7 ± 0.18	97.4 ± 0.24	94.7 ± 0.12	94.7 ± 0.14

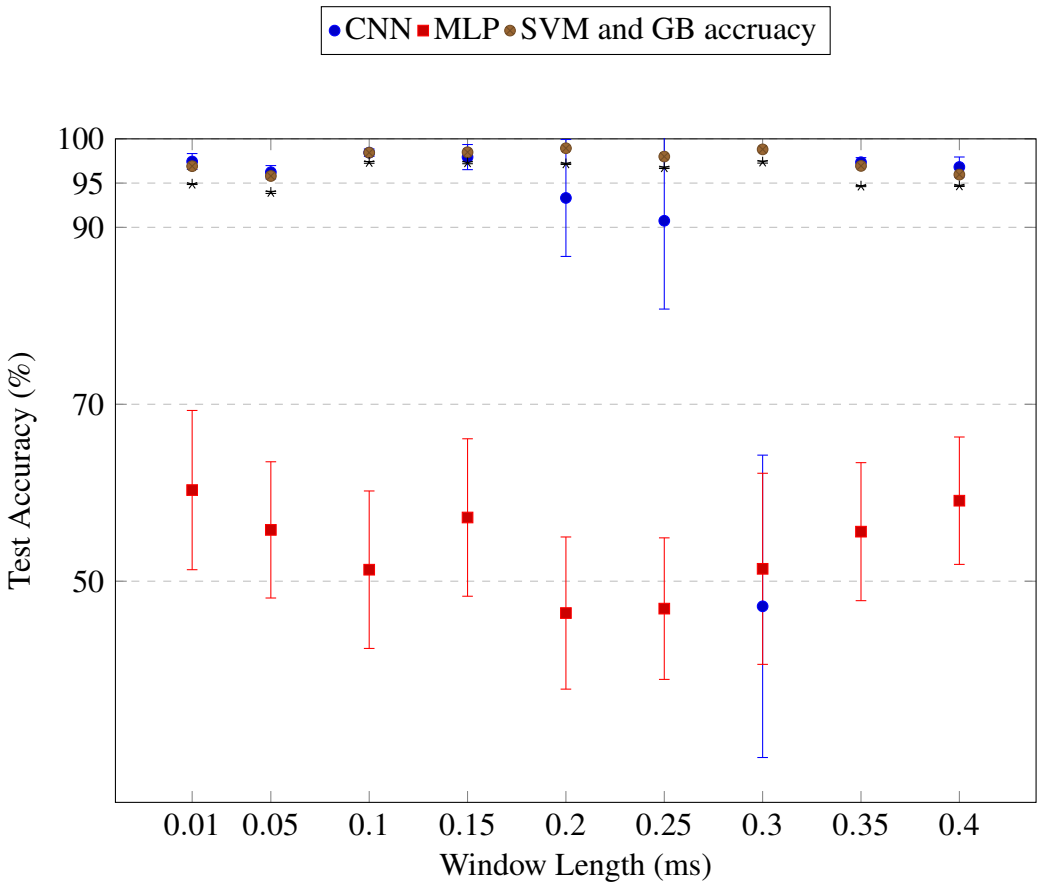


Figure 4.16. Performance of the CNN, MLP, SVM and GB classifiers for a range of window lengths L

4.2.5 Classifier comparison

The results presented in Table 4.3 and Figure 4.16 show that CNNs performs considerably better than MLPs for all variations of the window length L. MLPs performs poorly, with the highest accuracy $(60.3 \pm 18.1)\%$. Deviation of MLP results is also very high which can be interpreted as a very poor and unstable model fit to the data. SVM outperforms CNN for $L = 0.20$, $L = 0.25$ and $L = 0.30ms$ window length. This indicates that a SVM is able to generalize where a CNN cannot, specially for the case of $L = 0.30ms$. It can be a mild indication of overfitting, as a SVM is not as restricted as a CNN as how the input image is

interpreted and features extracted. A single pixel could be used by an SVM to produce a class decision. Gradient boosting shows good result for the classification task, but for most of the window length values it does not outperform both a CNN or SVM.

4.2.6 Window length comparison

The highest accuracy of (98.4 ± 0.8) % is achieved for window length $0.1ms$ for the CNN classifier. The results presented in Table 4.3 and Figure 4.16 show that $0.1ms$ window length provides a clear visual representation of the scattering with a fine compromise of time and frequency resolutions. SVM and GB present a stable results for different window length with the highest accuracy for gradient boosting for window length $0.15ms$. SVM performs with the best accuracy for window length $0.2ms$. MLP classifier has a very low accuracy and not considered in the comparison. It can be concluded, that a high accuracy classification can be achieved for a fine trade off the time and frequency resolutions. The CNN classifier obtained considerably lower performance at $L = 0.30ms$, the accuracy drop to $(47.2 \pm 34.2)\%$, but its performance increases after that point. This accuracy fall is appeared only for the CNN classifier. It is observed that by changing the STFT window length one can achieve high performance classification finding a compromise of time and frequency resolution based on the visual representation.

4.3 Filler based classification

In previous section, classes for classification are organised the way, that one class contains one object. In this section, objects are divided into two classes. The classes are formed based on the filler of the object: water and air. The objects are described by a form function. Wideband pulses are used to compute it. The form function is a signature of an object which depends on the object's properties including shape, size, material of the object's shell and material of the object's filler. The targets are classified based on the form function with the MLP neural network and support vector machine classifiers. There is a number of approaches which provide an insight for object classification based on the target scattering [47], [4], [63] and [41]. The novelty of this method is the combination of the form function and a neural network classifier for the filler material based classification. The target classification can be described by a three step process, Figure 4.17.

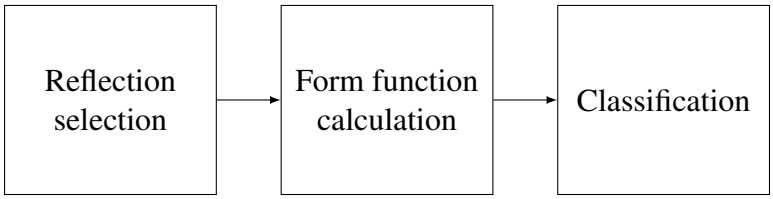


Figure 4.17. Scheme of the filler based classification

Firstly, the reflection from the object is selected from the recording. This process is

described in Section 4.1. The second step is calculation of the form function, as a descriptor of the target. In the third step, the form function is classified using MLP and SVM.

4.3.1 Form function

The form function expresses a pressure field scattered from a target in a range of frequencies [97]. It describes the way an object scatters a pulse, which makes the form function a good descriptor of an object itself. In this work, the form function is calculated from scattered and incident pulses with a knowledge about the medium where the signal was propagating. The frequency range of the function is limited by the bandwidth of the initial pulse. The form function is computed based on [47], where the scattered echo is calculated from a form function and initial pulse. Reformulating the task, the form function can be found from an experimental data by Equation 4.8.

$$f_\infty = \frac{|k_L^{(1)}|r^2}{e^{-2jrk_L^{(1)}}} \frac{FT[s(t)]}{FT[s_i(t)]}, \tag{4.8}$$

where r is a distance between source and the target, the value is calculated based on the matched filtering; $k_L^{(1)}$ is the wave number of the outer medium; $s(t)$ is a reflected pulse, $s_i(t)$ - initial pulse and FT is a Fourier transform, which is calculated using fast Fourier transform algorithm.

The distance to the object r is a parameter which has to be evaluated for each recording separately. The distance between the sonar's transducer and the object is computed based on the matched filter output, Figure 4.18. The distance r is proportional to the distance

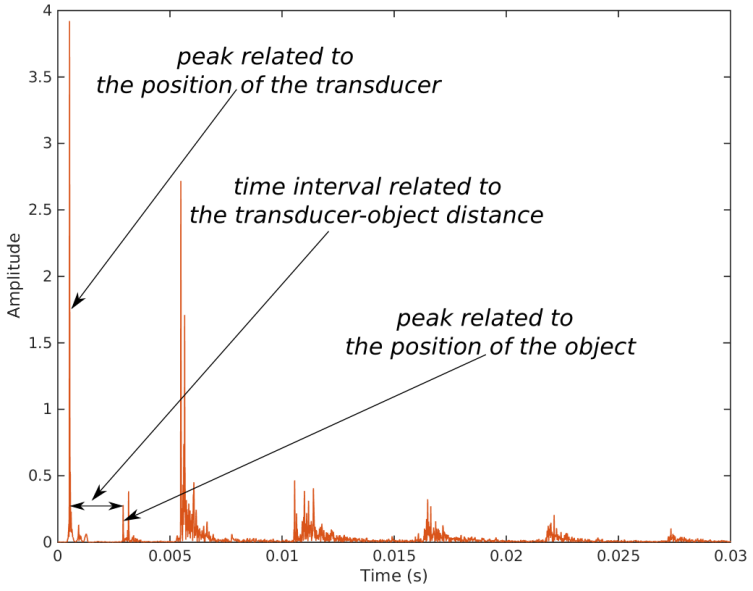


Figure 4.18. Distance between the peaks on the matched filter signal, which proportional to the distance between the sonar's transducer and the object: recording duration $0.03s$, sampling frequency $1MHz$, initial pulse duration $1ms$

between the first peak, which presents position of the transducer, and the peak detected as

the reflected from the target, Δt, Equation 4.9. The distance is highlighted in Figure 4.18.

$$r = \Delta t \frac{c}{2}, \tag{4.9}$$

where c is a speed of sound in the outer medium (i.e. the fresh water in the tank).

Figure 4.19 illustrates an example of the form function calculated for the reflection from an aluminium sphere filled with water (object 1 in Table 4.4). The form function is handled as an object descriptor and fed into machine learning classifier.

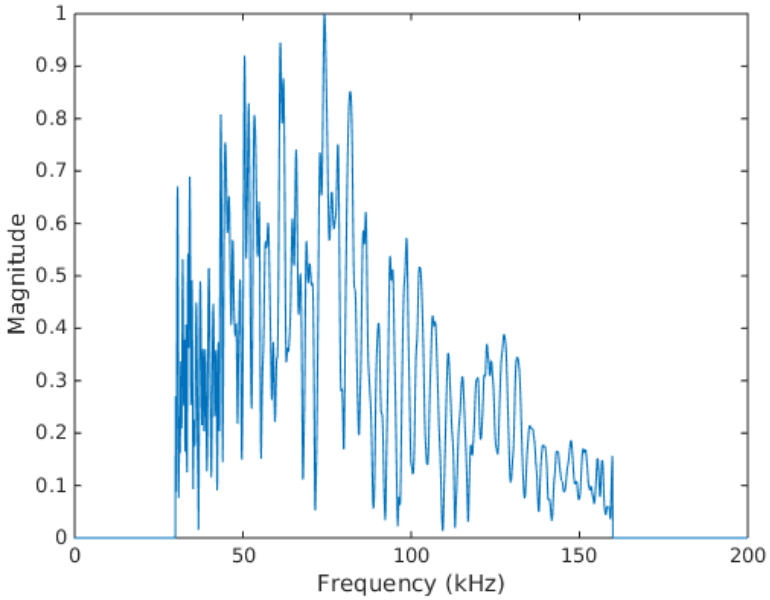


Figure 4.19. Obtained form function of an aluminium sphere filled with water

4.3.2 Classification

In this work, two classifiers were applied: a neural network classifier - multilayer perceptron (MLP) and support vector machine (SVM).

Multilayer perceptron

The general high performance of neural networks was a motivation for the choice of a neural network as a classifier. Even though multilayer perceptron (MLP) didn't perform well for the classification with TFD image in Section 4.2, it is chosen for the classification of the form function. Different more complex architecture of the MLP was chosen for the task, which allows generalise better 1D form function and provides high performance of the neural network. The network has a five fully connected layers configuration. The architecture is presented in Figure 4.20. The FC layers, except the output layer, are followed

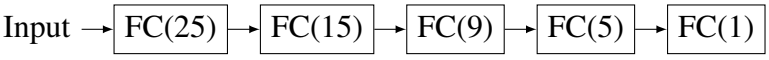


Figure 4.20. Multilayer perceptron model architecture

by dropout layers to prevent overfitting. ReLU activations (Equation 4.5) are used for all the layers except the output layer, which uses sigmoid activation (Equation 4.10).

$$f(x) = \frac{1}{1+e^{-x}} \tag{4.10}$$

Support vector machine

The support vector machine classifier is chosen with a radial basis function (RBF) kernel [107]. The classification is performed with 3-fold cross-validation for both classifiers to eliminate effect of data on the classification accuracy and provide fair evaluation of the results.

The described solution is implemented in Python.

4.3.3 Classification classes

There are 15 different objects classified into two classes: filled with air and filled with water. It is important to mention, that the same sphere, filled with different fillers, count as different objects. Parameters of the objects are listed in Table 4.4.

Table 4.4. Description of the objects

object	diameter, m	shell material	filler material
1	0.150	aluminium	water
2	0.130	stainless steel	air
3	0.130	stainless steel	water
4	0.200	stainless steel	air
5	0.200	stainless steel	water
6	0.280	stainless steel	air
7	0.280	stainless steel	water
8	0.120	plastic	water
9	0.120	plastic	air
10	0.200	abs material	water
11	0.200	abs material	air
12	0.275	plastic	water
13	0.275	plastic	air
14	0.210	plastic	water
15	0.210	plastic	air

The targets were insonified with a $1ms$ linear down-chirp pulse, which is described in Section 4.1. The recordings were made for the objects from different angles and distances. The obtained dataset contains 860 recordings with 430 examples per class.

4.3.4 Results

The obtained dataset is classified into 2 classes based on the filler material. Performance of the form function as a descriptor is compared to the echo representation in frequency

domain (FD). Both of the descriptors exist in the FD, but the form function is associated with the object itself, the FD representation expresses the reflected signal. Results for both MLP and SVM classifiers are presented in Table 4.5.

Table 4.5. Accuracy of the classifiers

Echo Representation	Accuracy (%)	
	MLP	SVM
Form function	98.60 ± 1.31	91.98 ± 1.78
Frequency domain	92.43 ± 3.41	52.67 ± 5.28

The highest accuracy is achieved with form function based MLP classifier. The results show that MLP classifier outperforms SVM. The form function based classification results present higher accuracy for both classifiers. The performance of the SVM classifier demonstrates significant advance in the accuracy for the form function case, while MLP results have less critical difference. The results present advantage of using the form function descriptor and illustrate that the form function is a stable feature vector which doesn't depend on the initial pulse and suitable for classification of object's materials. In this section, material based classification is implemented, where objects are classified based on their filler material.

4.4 Sea floor recognition with wideband pulses

Another application of the wideband pulses and machine learning techniques is a sea floor recognition. The task of the seabed mapping using sonar images is in a widespread use, nevertheless the recognition of some seabed surfaces can be complicated due to variation in their composition and roughness. The wideband pulses can be analysed in the frequency domain, where the difference in the seabed surface can be evaluated based on the frequency component distribution. This section presents the seabed recognition based on the frequency components of the reflected signal and compares different frequency based approaches.

Part of the sea floor of Tallinn bay was surveyed using a wideband sonar. The sonar supports sea floor study with chirp-based pulses over a wide range of frequencies ($30kHz - 160kHz$). It permits the possibility to study the frequency components of the reflected signals. This work considers a case when seabed was insonified with a single linear up-chirp pulses. The area of the Tallinn bay contains citadel constructions and sediment regions. The recognition task is to distinguish between these two types of the sea floor. It was observed that the frequency spectrum of the signals returned from these two surfaces has different distributions.

The work presents three frequency based approaches for the feature extraction process: time-frequency moment singular value decomposition (TFM-SVD), energy vector and complete frequency spectrum. Comparison of the techniques is made in terms of the classification results. Classification stage is implemented with support vector machine.

4.4.1 Data description

The data was recorded in Tallinn bay, where ruins of the 18th century artillery battery Citadel are located off the shore of Kadriorg, Tallinn, Estonia. The area is about 900m from the shore, at a depth of $(8-11)m$. The Citadel contains three separate standing cell boxes made of logs and filled with stones, 150×100 metres in total, Figure 4.21. The citadel plan is provided by Estonian Maritime Museum, Tallinn, Estonia.

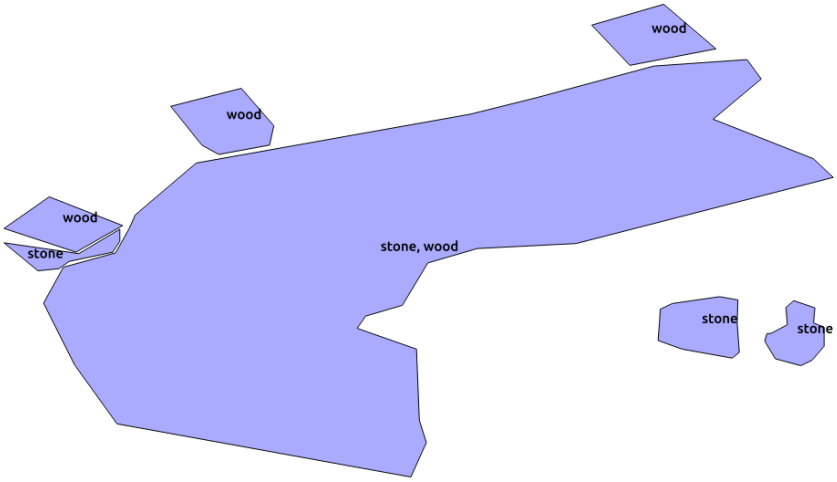
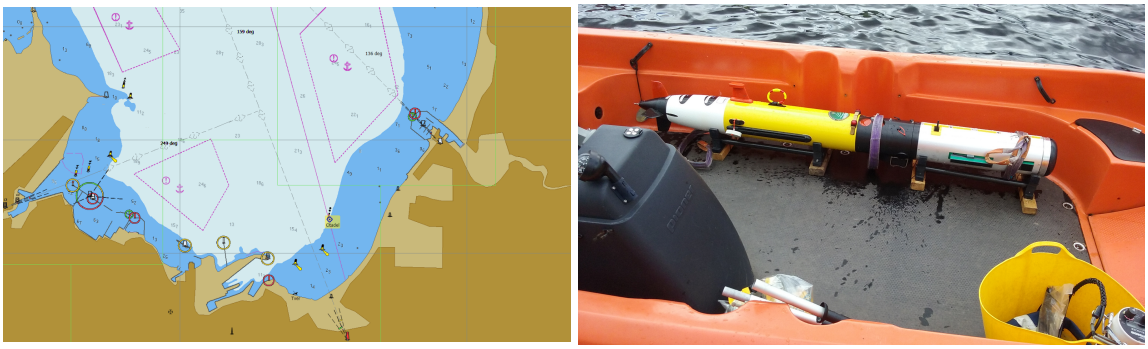


Figure 4.21. Plan of the citadel in Tallinn bay

The survey was made during the ARROWS FP7 project trials by a team of Heriot-Watt University and Hydrason Ltd. The wideband sonar was mounted on the Autonomous Underwater Vehicle (AUV) REMUS-100, Figure 4.22b. The vehicle had made four pro-



(a) Map of the Citadel

(b) REMUS with the Biosonar

Figure 4.22. Recording of the data

grammed missions (including compass calibration) in the area of the Citadel. Figure 4.23 presents one of the missions, which is projected into the citadel scheme using QGIS software.

The projection of the path is made based on coordinates provided by recordings of the REMUS navigation data during the mission. This mission took place above the citadel construction and the data from this particular mission is used in this work for the sea floor recognition task.

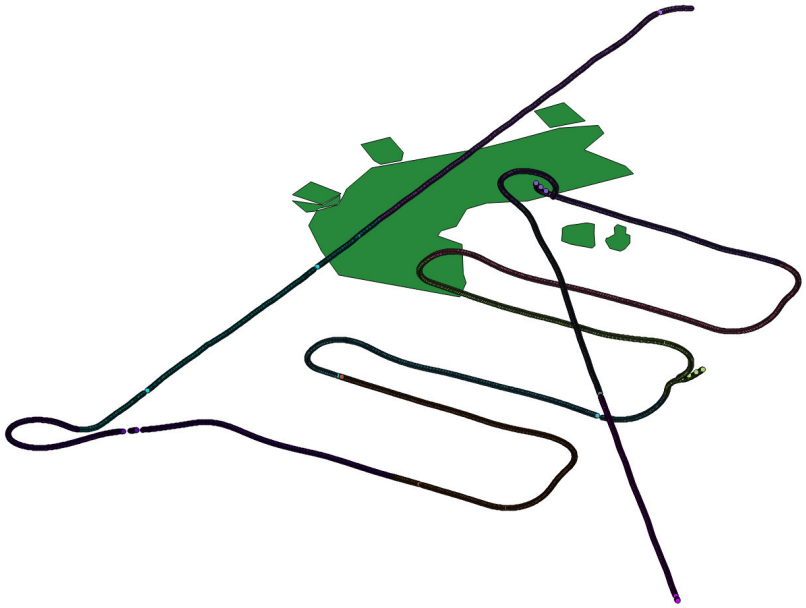
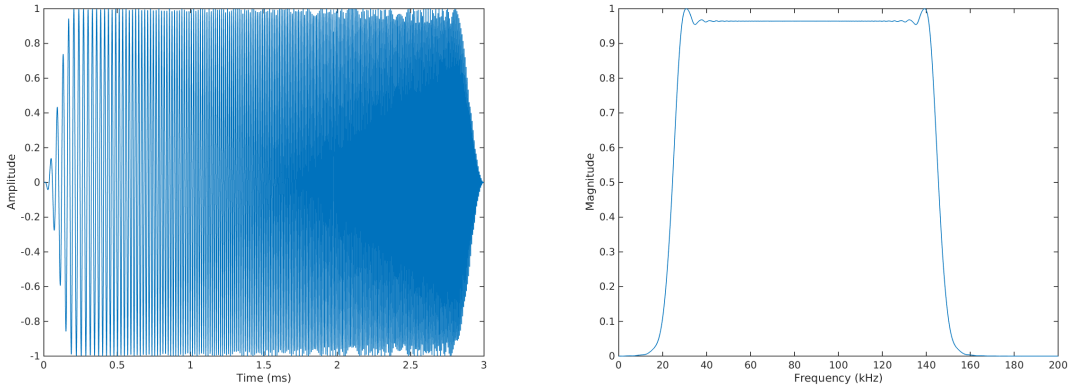


Figure 4.23. Remus mission projected into the citadel scheme

The biosonar is equipped with two transmitter-receiver pairs. The sonar works in the $30kHz$ to $160kHz$ frequency range and is able to transmit chirp-based pulses with this bandwidth. The seabed of Tallin bay was insonified by a single up-chirp pulse over the frequency range from $30kHz$ to $160kHz$ with a duration of $3ms$, Figure 4.24.



(a) Time representation

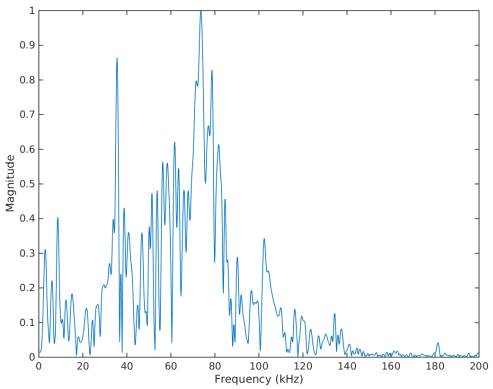
(b) Frequency representation

Figure 4.24. The transmitted up-chirp pulse: sampling frequency $1MHz$, pulse duration $3ms$

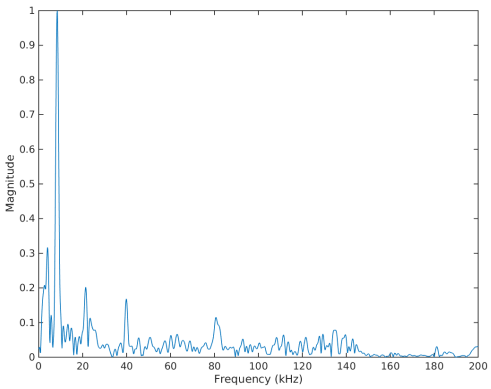
The biosonar doesn't have a connection to the REMUS's on-board computer. It uses power supply of the AUV, but the vehicle mission itself and biosonar recordings run separately. Therefore the wideband sonar records its data independently from the REMUS mission. Validation of the data is based on the REMUS sidescan sonar data and the AUV navigation data.

This work considers the data, based on knowledge about two types of sea floor surface. The two types to be recognised are sediment areas and citadel constructions. Figure 4.25 presents the one-sided spectrum of the returned pulses for both cases. Each class is presented by two examples. The examples show that the frequency spectrum of the reflected

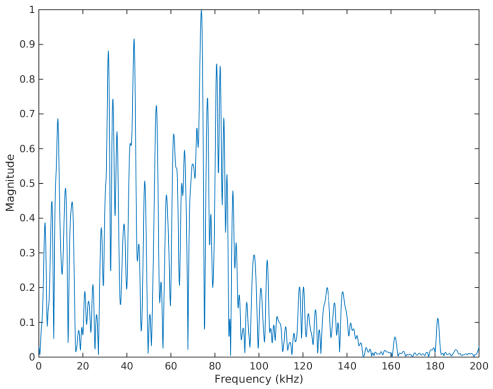
pulse from the citadel area can vary and not always look similar to each other, as well as for the sediment area. Nevertheless there is a noticeable difference between the sediment and citadel seabed types. The difference in the frequency components can be used for the recognition process. The example is representative, nevertheless there are more compli-



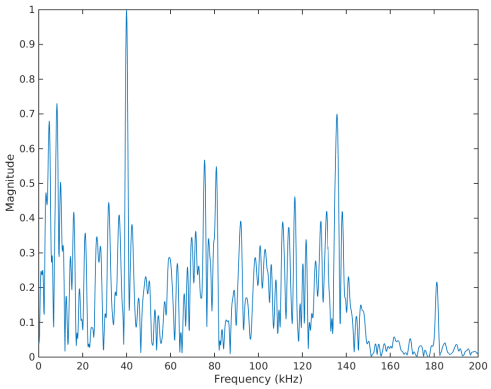
(a) Citadel: example 1



(b) Sediment: example 1



(c) Citadel: example 2



(d) Sediment: example 2

Figure 4.25. Normalised frequency spectrum of the areas of interest: sediment and citadel. Sampling frequency $1MHz$, 2000 FFT size

cated cases, when the responses from the sediment and the citadel have less significant difference. Not all the captured data was used for the classification. Examples of the data were chosen based on their location and how descriptive they are.

4.4.2 Sea floor recognition

The sea floor recognition process, as with any other recognition task, can be presented by the following steps: preprocessing, segmentation, feature extraction and classification. Preprocessing includes filtering of the raw data to reduce noise, and other processes to improve the data representation. The segmentation is a decomposition of data into separate parts, so each segment presents information about one type of seabed. Results of the segmentation can be represented as a pattern vector. The feature extraction process finds features that best characterise the data in the pattern vector for easy classification. The result is presented as a feature vector. The classification stage places the feature vector into a class. In this work,

there is no preprocessing step. This way, the recognition of the seabed can be presented as a four stage process, Figure 4.26. The input signal on the scheme is a single ping reflected from a seabed, while the output is a type of the seabed. There are two output classes: sediment and citadel.

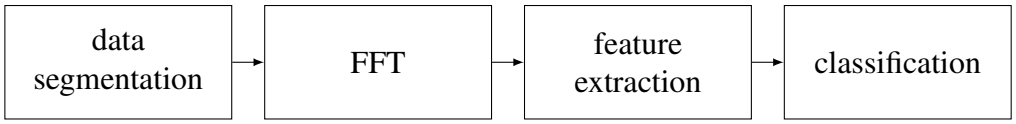


Figure 4.26. Scheme of the recognition process

At the data segmentation stage the returned ping is divided into smaller segments to process different parts of the seabed separately. This step is important for the frequency analysis, where the segmentation of the different surfaces is crucial. The Fourier transform step presents the segmented signal in frequency domain for feature extraction. The feature extraction stage is presented with three different approaches. Comparison of the approaches is the main objective of this work. The final classification step is based on the support vector machine classifier.

Segmentation

The data segmentation step breaks down each returned ping into smaller sections allowing separate evaluation of each section. Each segment is related to a small fragment of the sea floor area, Figure 4.27.

The length of the sector (window's size) is an important parameter for the final classification step. The smaller the sectors are the less information is fitted in one area and hence the separation of the citadel and sediment is more accurate. On the other hand, the length has to be long enough to provide sufficient resolution bandwidth (RBW) in the frequency domain. The resolution bandwidth can be calculated by the Equation 4.11, where f_s is a sampling frequency and N is a number of samples.

$$RBW = \frac{fs}{N} \tag{4.11}$$

Therefore the choice of the sector length is a compromise between amount of fitted information and the resolution bandwidth. The final segmentation was made with window length 1024, which provides 1 kHz resolution bandwidth. This setup brings necessary resolution to classify the sea floor based on the frequency spectrum.

There is a spectral leakage problem which is associated with finite observation interval. The segmented sectors are finite-time intervals. To reduce the leakage a weighting function can be applied to each segment. There is a large range of window types which can be used as a weighting function, including Gaussian, Blackman, Hamming, Blachman-Harris windows and others [103]. In this work, the Gaussian window was applied, as it is a non-negative, smooth, "bell-shaped" curve, and is a common solution for the task.

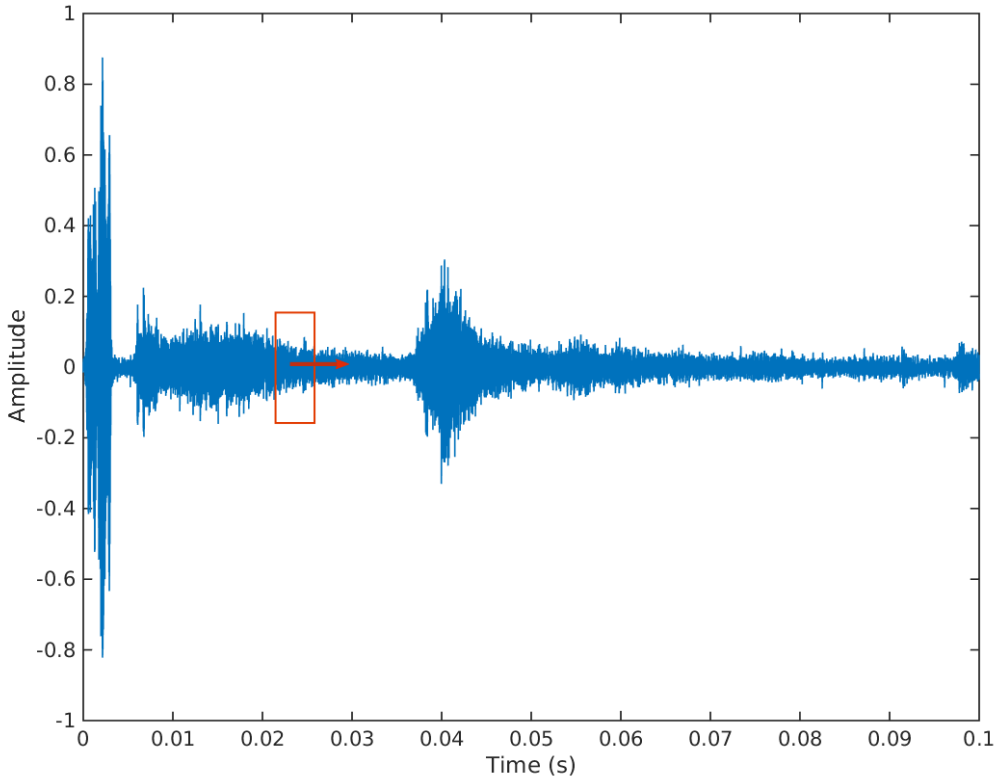


Figure 4.27. Returned single ping with a sliding window for segmentation: sampling frequency $1MHz$

Signal representation

The next stage presents each of the segments in frequency domain, using Discrete Fourier transform. For the analysis of the signals the first N/2 values are of interest, where N is the number of samples in the DFT. It forms the one-sided spectrum, Figure 4.28.

Feature extraction

The feature extraction stage finds the features that offer a way to characterise the data. Also, this step allows the reduction in dimensionality of the data. The work compares three different techniques: time-frequency moment singular value decomposition, energy vector and complete frequency spectrum.

Time-frequency moments singular value decomposition: The time-frequency moments singular value decomposition (TFM-SVD) method is introduced by Akhbardeh [108], as a part of a practical solution for ballistocardiogram data clustering. The proposed solution suits a periodic signal well. Hence the method considers only one period of the signal, it can be implemented for the task of biosonar response recognition.

The method is based on the calculation of statistical features of time and frequency series. The approach can be described in 4 steps: (1) compute statistical moments of time

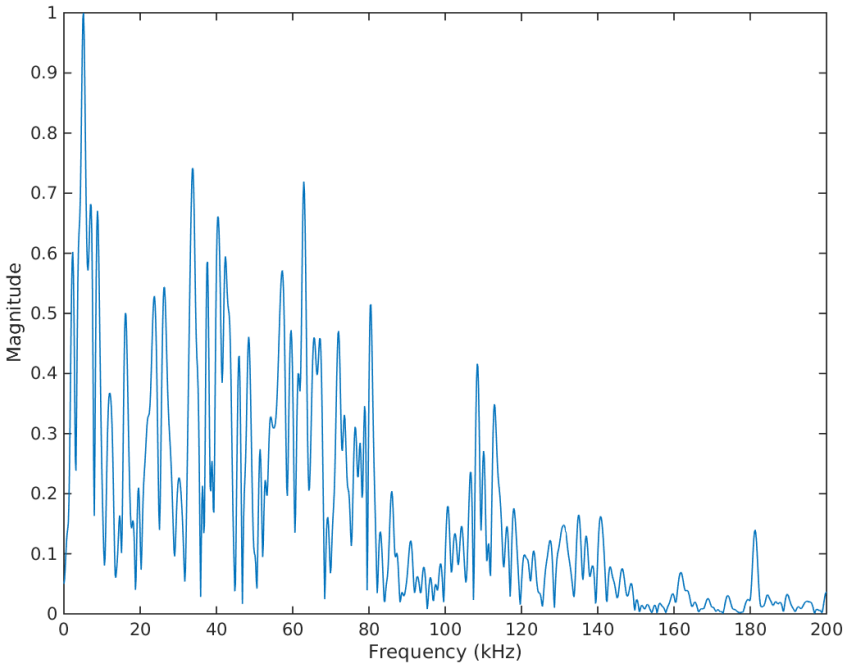


Figure 4.28. Normalised frequency spectrum of a segment: sampling frequency $1MHz$, 2000 FFT size

and frequency series, (2) assemble matrices which the calculates statistical moments, (3) calculate Singular Values of the matrices and (4) form the time-frequency moments vector.

There are four statistical moments, which represent mean value, variance, skewness and kurtosis, Equation 4.12.

$$\begin{aligned} mean &= E[x] = \mu \\ variance &= E[(x-\mu)^2] \\ skewness &= E[(\frac{x-\mu}{\sigma})^3] \\ kurtosis &= E[(\frac{x-\mu}{\sigma})^4] \end{aligned} \tag{4.12}$$

These statistical moments form two matrices M_t and M_f for time and frequency domain, Equation 4.13 and 4.14. The matrices combine statistical moments values to summarise the graph description.

$$M_t = \begin{bmatrix} a_t * mean_t & \frac{variance_t}{b_t} \\ \frac{skewness_t}{c_t} & \frac{kurtosis_t}{d_t} \end{bmatrix} \tag{4.13}$$

$$M_f = \begin{bmatrix} a_f * mean_f & \frac{variance_f}{b_f} \\ \frac{skewness_f}{c_f} & \frac{kurtosis_f}{d_f} \end{bmatrix} \tag{4.14}$$

Normalisation parameters a, b, c, d are used in the formula to regulate difference in the value orders. The choice of the coefficients for time and frequency matrices is described in [108], Equations 4.15 and 4.16.

$$\begin{aligned} a_t &= \frac{a_2 - a_1}{max(|a_1|, |a_2|)} \\ b_t &= a_2 - a_1 \\ c_t &= (a_2 - a_1)^2 \\ d_t &= (a_2 - a_1)^3 \end{aligned} \tag{4.15}$$

$$\begin{aligned} a_f &= \frac{a_2 - a_1}{N * max(|a_1|, |a_2|)} \\ b_f &= N^2 * (a_2 - a_1) \\ c_f &= N^3 * (a_2 - a_1)^2 \\ d_f &= N^4 * (a_2 - a_1)^3 \end{aligned} \tag{4.16}$$

The values of a_1 and a_2 are minimum and maximum values of the segment in time domain and N is the duration of the segment. Singular Value Decomposition (SVD) of the matrices defines features to describe the signal, Equation 4.17.

$$TFM = \begin{bmatrix} SVD(M_t) \\ SVD(M_f) \end{bmatrix} \tag{4.17}$$

Energy vector: The energy vector presents the frequency spectrum by a set of energy values for different frequency ranges. For this task energy needs to be calculated in chosen frequency ranges separately. Energy can be calculated from the discrete signal in time domain, $x[n]$, by Equation 4.18.

$$E = \sum_{n=0}^{N-1} |x[n]|^2 \tag{4.18}$$

Parseval's theorem can be applied here to evaluate energy using frequency spectrum.

$$E = \sum_{n=0}^{N-1} |x[n]|^2 = \frac{1}{N} \sum_{k=0}^{N-1} |X[k]|^2 \tag{4.19}$$

$|X[k]|^2$ is an energy density of the segment. Changing the sum's limits allows to calculate energy for different frequency ranges.

The approach allows the reduction of dimensionality, but reproducing the main character of the energy distribution in frequency spectrum. The spectral energy density of each segment (Figure 4.29) is divided into four equal frequency ranges: 30-60 kHz, 60-90 kHz, 90-120 kHz, 120-150 kHz. The number of frequency ranges is chosen based on the difference of the two spectrum types, considered for the classification. Energy is calculated for each frequency range separately. The values form an energy vector. The vector provides classifier with a distribution of the energy, but in a very low resolution scale.

Complete frequency spectrum: The complete frequency spectrum considers the spectrum itself in the frequency range of the emitted signal (from 30 kHz to 160 kHz). For the

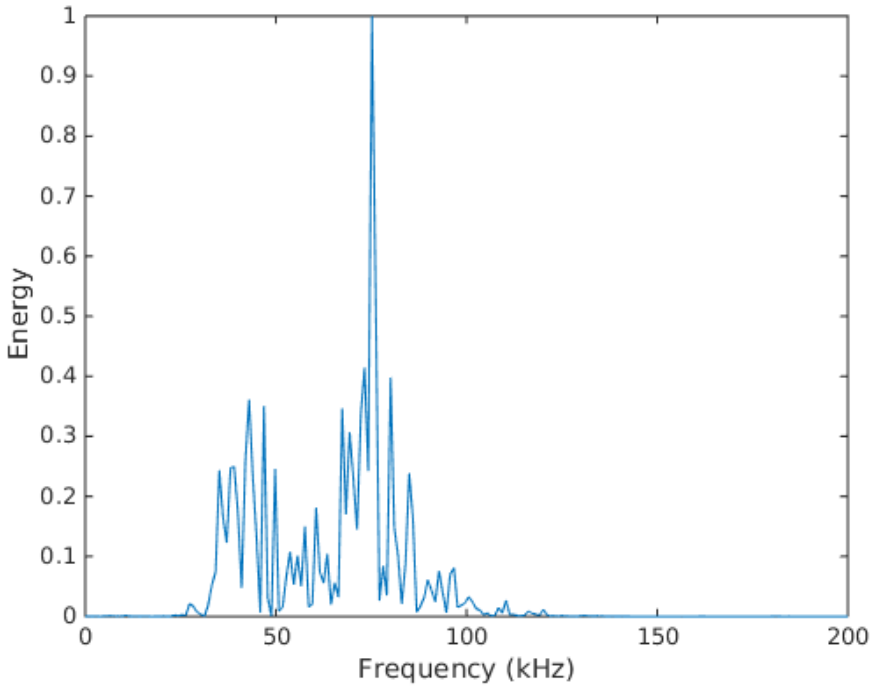


Figure 4.29. Normalised spectral energy density of a segment

analysis of the real valued signals, like the one considered in this work, the first half of the Fourier transform values are of interest. This approach doesn't reduce the data dimensionality, but provides all the details about the spectrum. It is presented for comparison with two other approaches.

Classification

Classification is presented by a supervised technique. There are two classes to be recognised: citadel and sediment. Binary support vector machine (SVM) approach was used in the work as a classifier with Gaussian Radial Basis Function kernel. The described solution is implemented in Matlab.

4.4.3 Results

The feature extraction techniques are compared based on the recognition results. The same SVM classifier was used for three feature extraction techniques. The approaches are compared by value of precision, recall and accuracy, presented in Table 4.6. Precision is defined as the number of true positives (elements correctly labelled as positive class) divided by the total number of elements labelled as positive class. Recall is the number of true positives divided by the total number of elements which belong to the positive class. Accuracy is calculated as a sum of true positives and true negatives (elements correctly labelled as negative class) divided by the number of all the elements of the testing set.

Precision highlights energy vector approach as the best one. Nevertheless the difference with two other approaches is not so big and the approach is evaluated very low by recall

Table 4.6. Statistical evaluation of the results

Approach	precision	recall	accuracy
TFM-SVD	0.85	1.00	0.91
Energy vector	0.86	0.65	0.77
Frequency spectrum	0.83	1.00	0.90

and accuracy. Recall is one for the TFM-SVD and frequency spectrum techniques. The accuracy result is a bit better for the TFM-SVD, than for the frequency spectrum, while the energy vector is a much poorer. Overall for the statistics evaluation, TFM-SVD approach shows the best results.

4.5 Summary

This chapter presented a machine learning based classification using wideband pulses. It showed how the variation in the echoes described in Chapter 3 can be used for classification of objects with different physical properties. Machine learning techniques were used to distinguish between objects with different characteristics. Firstly, objects with different physical properties were classified, where one class was represented by one sphere which can be considered as a combination of characteristics, such as radius, thickness, shell and filler materials. In the next step objects were classified based on only one characteristic - the filler material.

The objects that were used in this research were two-layered spherical shells. A dataset of wideband recordings was obtained for the classification. The experimental set up and initial pulse characteristics were provided. Influence of the sonar system on the reflected pulse was described by the transfer function. It was observed, that the high frequency region is influenced by the system with a visible damping after $85kHz$.

In Section 4.2, the object classification was demonstrated for a case where one object is represented by one class. The descriptor for the classifier is an energy distribution in wideband echo scattered from an object. The echo is represented in time-frequency domain (TFD), using short time Fourier transform (STFT) with different window lengths, and is fed into different classifiers, including convolution neural network (CNN), multilayer perceptron (MLP) classifier, support vector machine (SVM), and gradient boosting (GB). The results for different window lengths were analysed to study the influence of time and frequency resolution on classification. The CNN performed the best results with accuracy of $(98.44 \pm 0.8)\%$ over five object classes trained on grayscale TFD images with 0.1 ms window length of STFT. MLP classifier performed with a very low accuracy ($(60.3 \pm 0.04)\%$ is highest). SVM and GB presented a stable results for different window length with the highest accuracy for GB with window length 0.15ms. SVM performs with the best accuracy for window length 0.2ms. It can be concluded, that a high accuracy classification can be achieved for a fine trade off the time and frequency resolutions. This study showed good results for the classification of objects with the same shape and different physical properties.

The next step was a classification of objects based on one particular characteristic. Section 4.3 introduced object classification based on their filler material. Two classes of objects were formed for the study: filled with water and filled with air. The objects are described by their form function, which has distinguished peaks and notches related to the object properties. The form function is calculated from the reflected signal and handled as a descriptor for the classification with MLP and SVM. The results of the form function descriptor was compared to the representation of the echo in frequency domain. The highest accuracy of $(98.8 \pm 1.31)\%$ was achieved with form function descriptor and MLP classifier. The study illustrated possibility of object classification based on a single characteristic of the objects (filler material). In spite of the good results of the classification for the case of air and water, filler liquids with similar speeds of sound cannot be classified with high accuracy using this approach. The positioning of the peaks and notches varies less noticeably and other sphere's characteristic like size and shell material which have a bigger influence into the reflected pulse. To conclude, another more sensitive approach has to be developed to distinguish different materials. This approach is presented in Chapter 5.

In Section 4.4 of this chapter, wideband pulses and machine learning techniques were applied for the sea floor recognition task to demonstrate another possible application. The section provided comparison of three different descriptors: time-frequency moment singular value decomposition (TFM-SVD), energy vector and complete frequency spectrum. There were applied to the data recorded in Tallinn bay, Estonia, where the sea floor contains citadel constructions and sediment regions. TFM-SVD descriptor provided the highest accuracy of 91.0% and showed the best result for the sea floor recognition task among presented descriptors.

Chapter 5

Material identification based on the echo components

The true delight is in the finding out rather than in the knowing.

Isaac Asimov

In this chapter, an approach to an object's material identification is presented. The object is a metal sphere filled with liquid and placed in fresh water. The materials of the sphere's shell and filler are identified based on the timing of its echo's components. The echoes are formed by a number of processes which occur during the reflection of the initial pulse from the object. Reaching an object the sound wave reflects from the front and back walls of the object, propagates through the object and inside of the object's shell, as well as other paths, Figure 5.1.

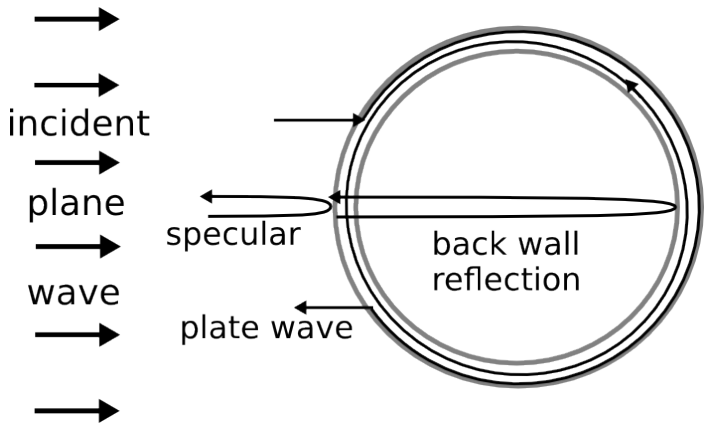


Figure 5.1. Some trajectories of the sound wave propagating in the sphere

The information about these components of the reflection can be extracted from the echoes and used to identify material. The wide bandwidth of the transmitted pulses is beneficial for the extraction of the reflection components' timing. A pulse compression approach [1] can be applied to the task, which is a practical implementation of matched filtering to a frequency-modulated signal. In this way, the components of the reflection can be determined using wideband chirp pulses. The timing of the reflection components and

geometry of the waves' paths determine speed of sound in the mediums where the wave is propagating. This calculated speed of sound is used to identify the material using a look-up table which provides speed in a number of expected materials. The sphere thickness is limited to 2% in relation to the sphere's radius. The results are evaluated for the sphere's radius in a range from $0.05m$ to $0.15m$. This material identification approach doesn't include any machine leaning techniques and doesn't require any training data. Material identification is based on the information about the sphere's radius, the outer medium and recording of the reflected echo.

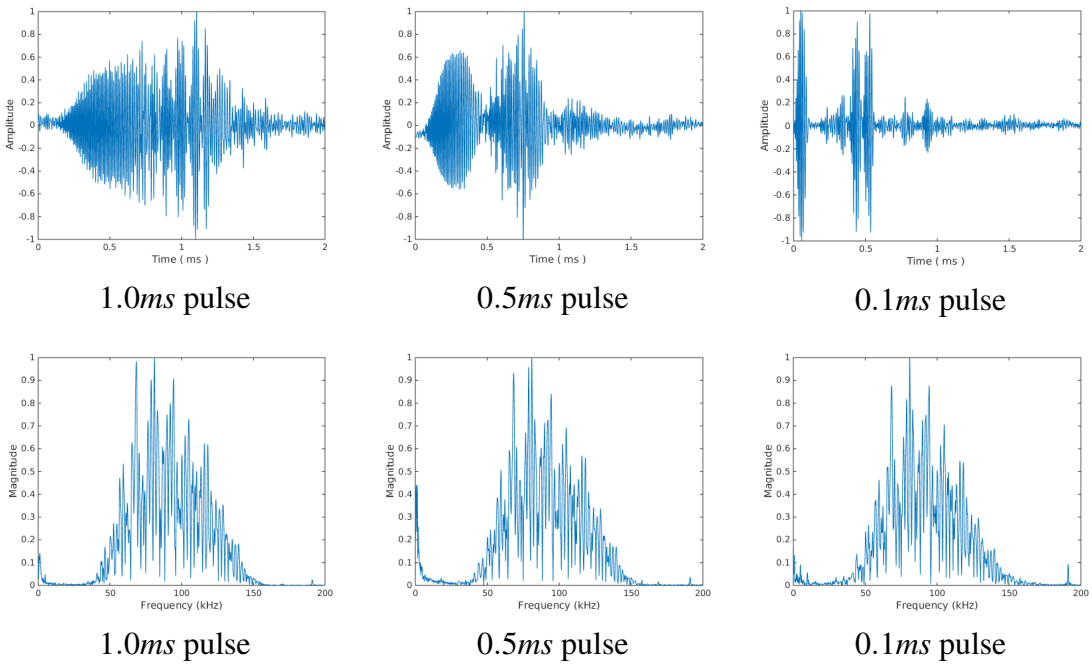
A description of the echo's components, which are used for the material identification is given in Section 5.1. This Section defines positioning of the components in the reflected pulse, their origin and presents a main concept for their timing estimation. Based on the components, the shell material identification approach is described in Section 5.2 and an approach for the filler identification is introduced in the Section 5.3. Results for both material identifiers are shown in Section 5.4. The results are demonstrated for experimental data, recorded in the test tank of the Ocean Systems Laboratory in Heriot-Watt University and for synthetic data generated by the approach described in Chapter 3.

5.1 Echo scattering

The echo reflected from an object usually has a different pulse shape and a frequency composition than the transmitted insonifying pulse. These changes in the pulse are due to the interaction of the sound wave with the object during the reflection process. The sound wave propagates by different trajectories inside of the sphere after reaching it. Each trajectory contributes an extra component to the final echo. In this section, an echo is considered as a composition of the elements produced by the different processes that are happening during the reflection.

The duration of the various reflected components depends on the length of the initial pulse. Each component of the reflected pulse is visible in time domain if the length of the initial pulse is short enough to ensure that there is no overlap. This is also dependent on the physical separation of the reflecting components and the type of wave that is produced. Figure 5.2 presents an example of the echoes recorded for three different durations of the initial pulse and the same object. It shows $2ms$ recording of the returned echo presented in time and frequency domains with sampling frequency $1MHz$ and FFT size 10001. The object is a stainless steel $0.1m$ radius sphere filled with a salt water and submerged in a fresh water.

The echo of $1.0ms$ pulse has all the components overlapped, their visible separation is not possible for this signal. Initial pulse with $0.5ms$ length produces echo with a few separable elements, but they still overlapping with each other - only primary and secondary parts of the reflection can be identified. Duration of $0.1ms$ provides with the best visibility of the reflection components in time domain. Duration of the initial pulse doesn't influence much the frequency composition of the pulse, but only timing and visibility of the reflection



1.0ms pulse 0.5ms pulse 0.1ms pulse

1.0ms pulse 0.5ms pulse 0.1ms pulse

Figure 5.2. Reflection of pulses with different current transmitted pulse lengths in the time domain (TD) - top row, and frequency domain (FD) - bottom row. Recording duration 2ms, sampling frequency 1MHz, FFT size 10001

components in time domain. The suitable pulse length depends on the sphere's size, smaller objects require shorter initial pulses.

5.1.1 Echo's components

The components of the echo are detected using a matched filtering technique, [109]. The matched filtering is obtained by convolving the returned echo, s_{echo}, with a conjugated time-reversed version of the initial pulse, $s_{initial}$. For the discrete signals, $s_{echo}[n]$ and $s_{initial}[n]$, the matched filtering is presented by Equation 5.1.

$$s_{mf}[k] = \sum_{n=-\infty}^{\infty} s_{echo}[n] s_{initial}[n-k] \tag{5.1}$$

The matched filter, as a part of the pulse compression sonar, is commonly used to increase the range resolution as well as the signal-to-noise ratio (SNR) in a noisy environment. In this work, the matched filter is applied to detect components of the reflection. The reflection components are highly coherent with the initial pulse, which is indicated by large values of the s_{mf} signal. The peaks of similarity formed by the large values highlight the echo's components, for instance, waves propagate by different trajectories inside of the sphere. In this way, the components of the echo are presented in terms of peaks on the matched filtered signal, Figure 5.3.

The matched filtering provides information regarding the presence of the reflection's components and their timing. This technique doesn't limit the duration of the initial pulse to a short pulse. However, if the detected component has to be extracted from the echo,

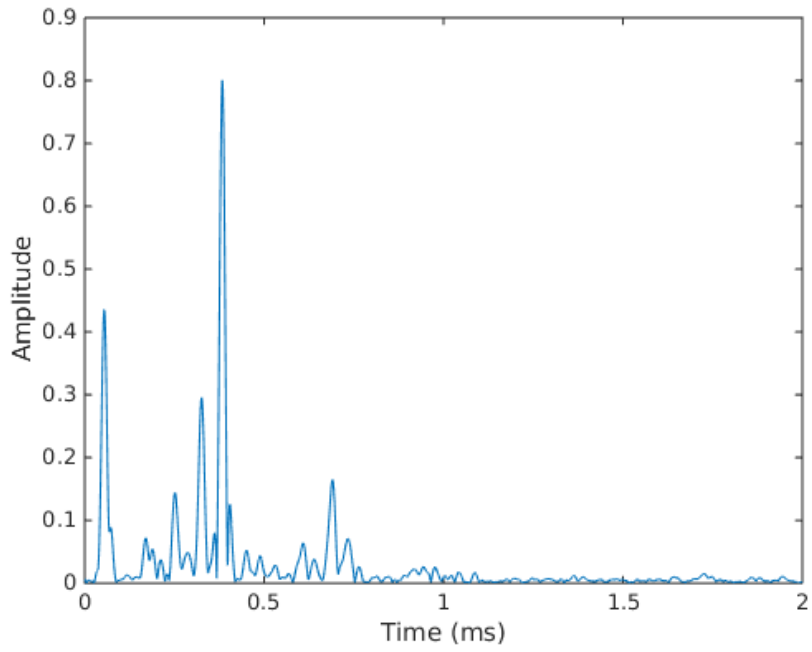


Figure 5.3. Matched filtering of an echo reflected from a stainless steel sphere filled with water and initial $0.1ms$ linear chirp pulse

the duration of the initial pulse has to be short enough to eliminate overlapping. Figure 5.4 demonstrates reflected echo with matched filter results and indicates echo components, which will be exploited in the material identification approach.

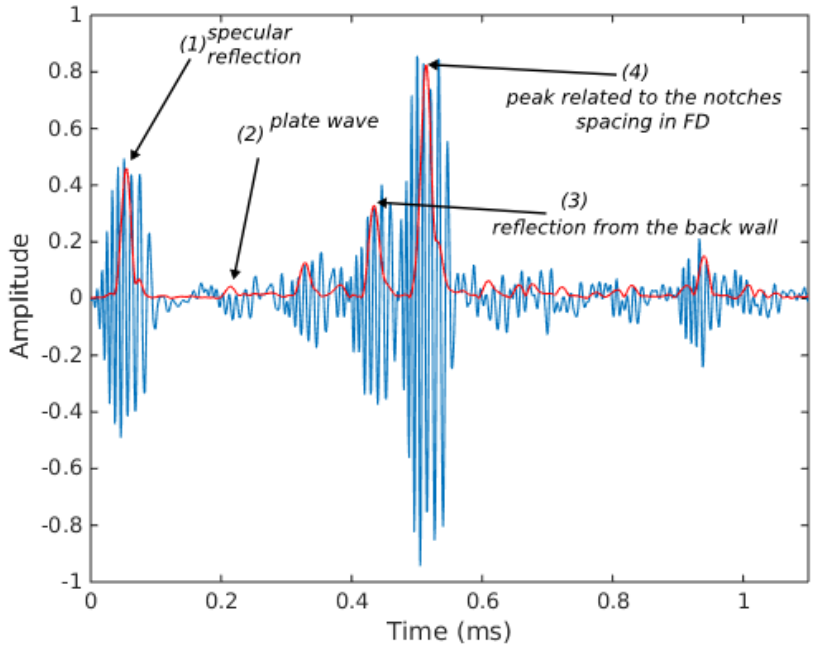


Figure 5.4. Components of the reflected echo: recorded echo in time domain (blue) and matched filter output(red)

The initial pulse is a $0.1ms$ down-chirp with 3 dB bandwidth of $(136-52)kHz$. The pulse is reflected from a stainless steel $0.14m$ radius sphere filled with fresh water and

placed in a water tank. The duration of the initial pulse and relatively large sphere size allow to identify components of the echo separately in the time domain. It is used for the demonstration purposes, while for the material identification method longer initial pulses are used.

The returned echo is composed by various processes: specular reflection, diffracted circumferential waves, elastic waves transmitted into the shell and bouncing from the inner shell surface, Lamb waves, etc. [4]. This section is focused on the components which are useful for the material identification. There are two major parts of the reflected pulse - the primary and secondary reflections. The primary echo is a specular reflection, which is a direct reflection from the front face of the sphere. The secondary reflection contains all the components which are propagated inside of the sphere and reflected back to the sonar. All the secondary components arrive with a delay after the primary echo.

Specular reflection (echo component 1 on the Figure 5.4) has a duration similar to the initial pulse. Figure 5.5 presents the specular component of the reflection and the initial pulse in time and frequency domains. It can be seen that the composition of the specular reflection similar to the initial pulse, the changes are due to the influence of the sonar's hardware and a noise during the transmission. The specular reflection doesn't contain in-

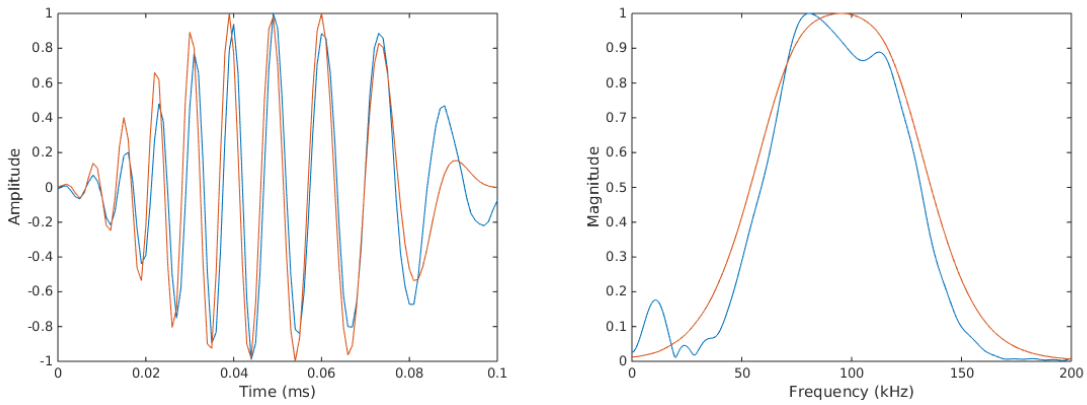


Figure 5.5. Specular reflection (blue) and transmitted pulse (red)

formation about the sphere's properties, for example sphere's size, shell's material or filler material. Nevertheless it can be used to evaluate distance from the sonar to the object , r, Equation 5.2.

$$r = \frac{c_1}{2t_{specular}}, \tag{5.2}$$

where c_1 is a speed of sound in the outer medium and $t_{specular}$ is a timing of the specular component. Moreover the timing of the specular reflection in combination with other measurements is used to calculate duration of the echo's components in the proposed material identification approach.

The secondary reflection contains a number of components, including Lamb-type waves. The Lamb-type waves are shell-borne surface waves propagating along the sphere in the shell's material. There are zero-order symmetric and asymmetrical Lamb-type waves propagating in a thin spherical shall in the given frequency range. Energy of the asymmetric

zero-order Lamb-type waves, a_0, is dumped very fast outside the shell and its contribution to echo is minimal. The symmetrical Lamb-type wave of the zero-order, s_0, contribute the most into the reflected pulse [63]. The plate wave (echo component 2 on the Figure 5.4) is a Lamb wave revolving around the shell in the material and leaking back into the outer medium. The wave propagates inside of the shell with a plate speed c_{pl} [4]. The plate wave is also called a Lamb wave group velocity [32]. The plate velocity can be calculated through Poisson's ratio ν:

$$c_{pl} = \frac{\sqrt{1-2\nu}}{1-\nu} c_L \tag{5.3}$$

$$\nu = \frac{c_L^2 - c_T^2}{2(c_L^2 - c_T^2)} \tag{5.4}$$

where c_L is a longitudinal speed of sound and c_T is a transversal speed of sound in the shell material. The Lamb wave velocity depends on the shell material. Therefore the timing of the plate wave is used for the shell material identification.

Another secondary reflection component related to the Lamb-type waves is peak 4 on the Figure 5.4. It is associated to the spacing between notches in frequency representation of the echo. The notches in frequency domain are formed mainly by the symmetric and asymmetric Lamb-wave of zero-order [4]. Example of the notches is presented on the Figure 5.6.

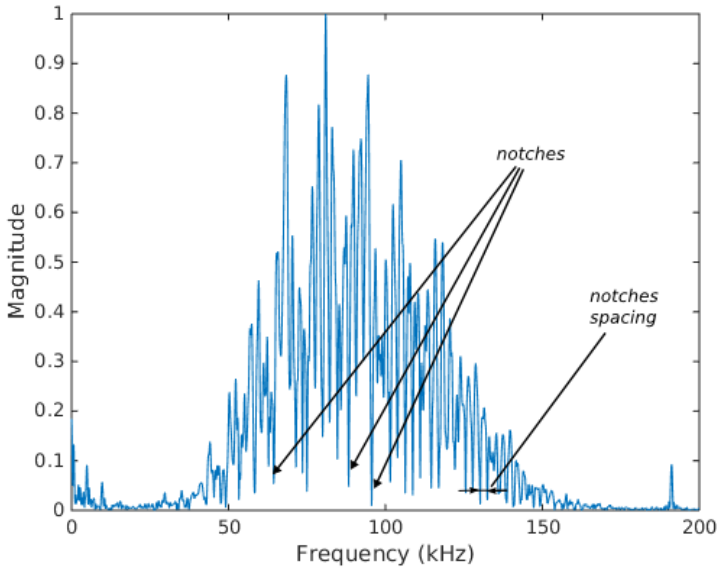


Figure 5.6. Notches and notches spacing in frequency domain of the returned echo. Recording duration - $2ms$, sampling frequency - $1MHz$, FFT size - 10001

The spacing between these notches is nearly constant and inverse from the timing of this peak. For the presented example of the $0.14m$ radius stainless steel sphere, the average value of the spacing between the notches, $\Delta f_{notches} = 2.15kHz$, while the timing of the peak 4 $\Delta t_{notches} = 0.459ms$ and $\Delta f_{notches} \times \Delta t_{notches} = 0.986$. The calculations were performed for all the experimental data and give similar results.

Peak 3 on the Figure 5.4 represents a wave which propagates through the shell, the filler and reflects from the back wall of the sphere. Timing of the wave depends on the speed of sound in the shell and in the filler. The timing of this elastic wave is used for the identification of the filler material.

It is worth mentioning one more component which belongs to the secondary reflection - a mid frequency enhancement (MFE). The enhancement is caused by the subsonic asymmetric Lamb wave, a_{0-}. Li *et al.* [41] demonstrates that elastic echo associated with the subsonic asymmetric Lamb wave becomes stronger than the specular reflection in the mid frequency range. Frequency of the MFE is proportional to the ratio of a/d. In this work, the MFE is not used for the sphere's parameter identification, but has to be highlighted as a component, which contains useful information.

It is important to mention, that the sound wave propagates by many different paths. This section covers only components which contribute the most into the echo and are essential for the material identification.

5.1.2 Timing of the components

As was mentioned before, the matched filtering allows calculation of the timing of the reflection's components. The time stamp when a component of the reflection occurred, t_{pos}, can be calculated from the timing of the peak, t_{comp}, and the initial pulse duration , Δt_i:

$$t_{pos} = t_{comp} - \Delta t_i/2 \tag{5.5}$$

The timing of the component is the duration of the sound wave travelling by its path. It can be calculated as a difference between the specular reflection and the moment of the reflection component when the wave came back to the point of the specular reflection, $\Delta t_{comp} = t_{pos} - t_{specular}$. Combining it with the Equation 5.5, the timing of the process can be calculated as a difference between the peaks:

$$\Delta t_{comp} = t_{comp} - t_{specular_{peak}} \tag{5.6}$$

The timing calculation presented in this section is used in the material identification methods.

5.1.3 Sphere's thickness limitation

The distribution of the echo's components depends on the thickness of the sphere. The example in Figure 5.4 presents reflection from a sphere with a thickness 0.7% of the radius. It gives a clear picture of the reflection, which allows extraction of the timing of the components and identify material, while the thickness increases more components of the secondary reflection appear. An example in Figure 5.7 presents $0.3ms$ of the echo's secondary components, which are influenced by the sphere's thickness. These examples are

analytically calculated echoes reflected from an aluminium sphere with radius $0.11m$ and thickness in a range from 1% to 5%.

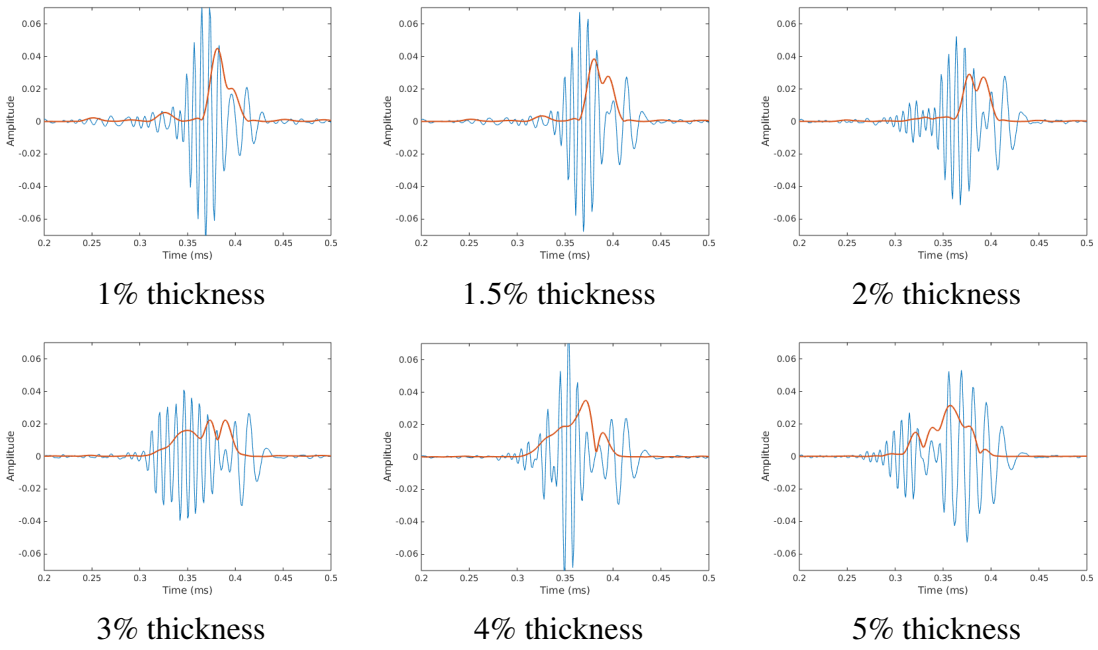


Figure 5.7. Part of the echo reflected from spheres with different thickness: recorded echo in time domain (blue) and matched filter output (red). The data is analytically calculated for an aluminium sphere with radius $0.11m$

Influence of the shell thickness can be examined based on how recognisable the reflection components in the matched filter signal. Thickness 1% and 1.5% provide a clear picture similar to the example in Figure 5.4, while thickness 2% has some extra components, which can cause a problem for the peak detection, but the reflection components are still recognisable. Reflections from a sphere with thickness higher than 2% have too many extra components and too many peaks in matched filter output. These peaks make it difficult to detect essential components for the material identification - the plate wave and the wave reflected from the back wall. Therefore for the presented approach of the material identification, the sphere's thickness d is limited to $d \leq 2\%$ of the radius.

5.1.4 Frequency range limitation

To be reflected by an object, the pulse wave length has to be related with the size of the object. In case of the spherical object, the wave lengths, λ, of the initial pulse must be adjusted to the radius of the sphere, a. It has to satisfy the limit: $\lambda < 2a$. Smaller λ provides with a more precise acquisition of the sphere. Therefore, the frequency range of the initial pulse has to be adopted to the object's size.

The frequency range, which is used in the work is $(30-160)kHz$. The lowest frequency of the band defines the minimum sphere's radius. For the given signal, the radius is restricted by $0.025m$. The calculations are made for the wave propagating in a fresh water ($1480m/s$). The smallest object's radius in the presented work is $0.05m$ for the simulated

data and $0.065m$ for the synthetic data. The frequency range is suitable for all set of objects used in this work.

5.2 Shell material evaluation

The material of the shell can be defined based on the timing of the plate wave, revolving around the sphere. The time delay between the specular reflection and the plate wave component provides information about the speed of sound in the shell material, which is used for the material identification. The approach can be described in five steps, Figure 5.8.

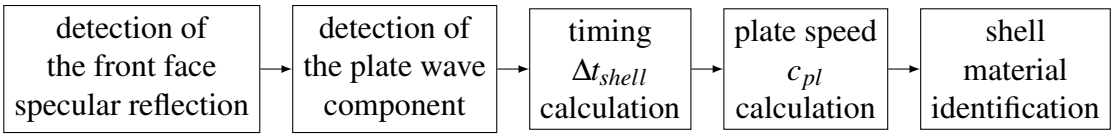


Figure 5.8. Scheme of the shell material identification

Firstly, the specular and plate wave components are detected based on the matched filtering of the reflected echo with the initial pulse. The timing of the plate wave is evaluated based on the detected components. The calculated speed of sound is used for the material identification, where the material of the shell is defined based on the closest speed value in a look-up table.

5.2.1 Plate wave propagation

The plate wave, a Lamb-type wave, can also be referred to as a surface elastic wave. It is important to notice that the displacements associated with Lamb waves are not necessarily placed near the surface of the plate [55], though for the simplification of the calculation we will consider the path of the Lamb wave on the surface of the shell. The wave is excited in the shell near the critical angle, θ_c, Figure 5.9. The critical angle depends on the speed of sound in the outer layer, c_1 and the Lamb wave group velocity in the shell, c_{pl} [32], Equation 5.7.

$$\theta_c = arcsin(\frac{c_1}{c_{pl}}) \tag{5.7}$$

The wave propagates in the solid material along the curved surface of the shell and radiated back into the outer medium near the angle $-\theta_c$. The path of the plate wave is presented in Figure 5.9, which includes some segments of the path in the outer medium. These segments are not a plate wave, but they have to be considered, as they contribute to the delay between specular and the plate wave components.

This trajectory can be divided into three parts. The first one is a path in the outer medium when the wave is reaching the shell, $a(1-cos\theta_c)$. The second part is a way along the curve surface of the shell, $2a(\pi-\theta_c)$. And the last one is in the outer medium back to the point aligned with the outer surface of the sphere, $a(1-cos\theta_c)$. The time duration of a sound

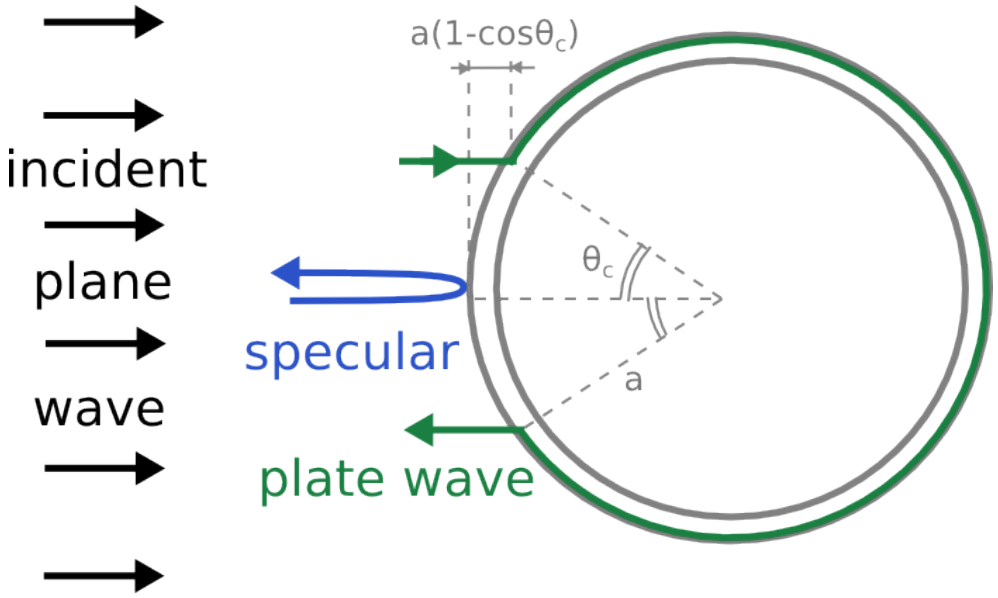


Figure 5.9. Trajectory of the plate wave component

wave propagating along the path can provide an insight into the speed of sound in the shell material.

The sound wave is travelling outside and inside the sphere, the timing can be calculated as a sum of two components. The first component is a time difference between the specular reflection and the wavefront reaching the shell at a critical angle. The component is counted twice as the same time is needed for the wave to come back to the same range as the specular reflection. At this step the wave is propagating in the outer medium (water). The timing is related to the sphere radius, outer medium and material of the sphere's shell. The second element is a timing of the wave propagating around the shell and diffracted at the angle $-\theta_c$. It is the plate wave trajectory along the sphere's shell. Duration of the second component depends completely on the sphere radius and material of the shell. The timing of the plate wave, Δt_{shell}, can be calculated by Equation 5.8 [32].

$$\Delta t_{shell} = 2a\left(\frac{1-cos\theta_c}{c_1} + \frac{\pi-\theta_c}{c_{pl}}\right), \tag{5.8}$$

where a is a radius of the sphere.

The outer medium is fresh water, speed of sound of the water in the calculations is $1480 m/s$. The radius of the sphere and material of the shell are varied. Table 5.1 presents timing of the plate wave revolving around the sphere with different radius and shell materials.

This table illustrates difference in timing of the plate wave propagation. The radius and speed of sound in the shell material influences the timing. The properties of the shell materials are presented in Table 5.2. The timing Δt_{shell} increases with radius and decreases when the plate speed of the shell increases.

Based on the Equation 5.8, the speed of sound in the shell material can be defined from the Δt_{shell} measurement with a known values of the sphere's radius and speed of sound in

Table 5.1. Timing of the wave propagating through the sphere's shell with different radius and shell materials

	Δt_{shell} (μs)						
sphere radius	$0.05m$	$0.075m$	$0.1m$	$0.125m$	$0.15m$	$0.175m$	$0.2m$
aluminium	55	83	110	128	166	193	221
stainless steel 347	57	86	114	143	172	200	229
copper	73	110	147	183	220	256	293
brass	76	114	152	190	228	266	302

Table 5.2. Properties of the shell materials

Materials	Density kg/m^3	Long. speed m/s	Trans. speed m/s	Plate speed m/s
aluminium	2700	6320	3130	5438
stainless steel 347	7910	5740	3090	5208
copper	8930	4660	2330	4035
brass	8560	4280	2300	3879

the outer medium.

5.2.2 Timing of the plate wave

The plate wave timing Δt_{shell} can be measured as a time delay between specular reflection and the plate wave, Equation 5.9.

$$\Delta t_{shell} = t_{plate} - t_{specular} \tag{5.9}$$

The exact timing of the both specular and the plate wave components can be detected using matched filtering technique. First, the specular reflection is found as the first large peak in a given time interval, Algorithm 2.

Input of the algorithm is a radius a, speed of sound in the outer medium , c_1, reflected echo in time domain, $S_{echo}(t)$ and coefficient k_{mmp}. The matched filtering output $s_{mf}(t)$ is searched for peaks in a range from the beginning of the echo pulse until $t_{\frac{a}{2}}$. The time $t_{\frac{a}{2}}$ is a duration which sound wave takes to propagate in the outer medium for a half a radius distance. This time was chosen experimentally, in this range the first large peak always will be the specular reflection.

The peak detector identifies peaks as a data points which are larger than its two neighbouring samples. The endpoints of the signal $s_{mf}(t)$ are excluded from the detection. Only the peaks with a prominence higher than the minimum peak prominence (MPP) value are taken into consideration. The peak's values, P_{val}, and locations , P_{loc}, are presented in a form of arrays. The arrays are arranged in the order of the peak positioning on the time axis. The coefficient k_{mpp} defines a ratio of MPP in relation to the matched filtering signals values: $MPP = k_{mpp} * max(s_{mf}(t))$.

Algorithm 2 Timing of the specular reflection

Input a, c_1, k_{mmp}, $s_{echo}(t)$, $s_i(t)$

Output $t_{specular}$

matched filtering $s_{mf}(t)$
calculate $t_{\frac{a}{2}} = \frac{a}{2c_1}$
calculate $MPP = k_{MPP} * max(s_{mf}(t))$
find peaks P_{loc} and P_{val} in $s_{mf}(0 : t_{\frac{a}{2}})$
while $length(P_{loc}) < 1$ & $MPP >= k_{MPP} * max(s_{mf}(t)) * 0.01$ **do**
    $MPP = MPP - 0.01MPP$
    find peaks P_{loc} and P_{val} in $s_{mf}(0 : t_{\frac{a}{2}})$
    if $MPP < k_{MPP} * max(s_{mf}(t)) * 0.01$ **then**
        position of the largest peak $N_{plate} = 0$
        break
    end
end
if $length(P_{loc}) >= 1$ **then**
    peak related to the specular reflection: $N_{specular} = 1$
    calculate $t_{specular} = P_{pos}(N_{specular})$
else
    $t_{specular} = length(s_i(t))/2$
end

The prominence of the specular peak depends on the sphere's parameters. Therefore MPP value is tuned based on the peak detector outcome (*while loop*). If no peaks are detected, then the values of MPP is decreased by 1% until at least one peak is found or the MPP becomes too small. If the peak is not found, then it is assumed that the specular reflection locates at the beginning of the echo and it's timing is calculated as $t_{specular} = length(s_i(t))/2$. This approach provides adaptive tuning of the peak detector based on the matched filtering output. The solution is implemented in Matlab, as well as for all the following algorithms.

The plate wave component is the first large secondary reflection after the specular echo. Therefore the peak corresponding to the plate wave component is searched for after the specular reflection with a limit at t_{2a}, the time duration which a sound wave needs to propagate in the outer medium for the two radii distance. These boundaries eliminate detection of the secondary reflections from the back wall of the shell and other large secondary components, which have a larger time delay. The moment of the maximum peak in the range is marked as the timing of the plate wave, t_{plate}. The method is described by Algorithm 3.

Algorithm 3 Timing of the plate wave component

Input a, c_1, k_{mmp}, $s_{echo}(t)$, $t_{specular}$, $s_i(t)$

Output t_{plate}

matched filtering $s_{mf}(t)$
calculate $t_{2a} = \frac{2a}{c_1}$
calculate $MPP = k_{MPP} * max(s_{mf}(t))$
find peaks P_{loc} and P_{val} in $s_{mf}(t_{specular} : t_{2a})$
if $length(P_{loc}) <= 1$ **then**
    while $length(P_{loc}) < 1$ & $MPP >= k_{MPP} * max(s_{mf}(t)) * 0.01$ **do**
        $MPP = MPP - 0.01MPP$
        find peaks P_{loc} and P_{val} in $s_{mf}(t_{specular} : t_{2a})$
        if $MPP < k_{MPP} * max(s_{mf}(t)) * 0.01$ **then**
            position of the largest peak $N_{plate} = 0$
            break
        end
    end
    if $length(P_{loc}) > 1$ **then**
        find the peak related to the plate wave: $N_{plate} = max(P_{val})$
    end
else
    find the peak related to the plate wave: $N_{plate} = max(P_{val})$
end
if $N_{plate} == 0$ **then**
    print " t_{plate} is not found"
else
    calculate $t_{plate} = P_{pos}(N_{plate}) + t_{specular}$
end

This algorithm defines t_{plate} from the radius a, speed of sound in the outer medium, c_1, reflected echo in time domain, $S_{echo}(t)$, coefficient k_{mmp} and time stamp of specular reflection, $t_{specular}$. Coefficient k_{mmp} influences peak detector sensitivity for the prominence of the peaks in the matched filtering signal. The peak detector considers peaks only with a prominence higher than the minimum peak prominence. MPP value is tuned based on the peak detector outcome the same way as it is implemented for the specular component detector. If there is one or less detected peak, then the value of MPP is decreased by 1% until more peaks are found or the MPP values becomes too small. In the case where $MPP >= k_{MPP} * max(s_{mf}(t)) * 0.01$ the peak detector is stopped and t_{plate} value cannot be detected. The largest detected peak is marked as a plate wave component and its value is calculated as a sum of the peak position and the beginning of the peak detection range: $t_{plate} = P_{pos}(N_{plate}) + t_{specular}$.

The final plate wave timing Δt_{shell} can be calculated based on the Equation 5.9. This value is used to define speed of sound in the shell material.

5.2.3 Speed of sound in the material

The speed of sound in the shell material and timing of the plate wave component is related through the Equation 5.8. The speed of sound in an outer layer, c_1, is assumed to be $1480m/s$, as it is a fresh water for all the cases. The critical angle θ_c depends on the plate speed of the shell material and speed of sound in the water. Algorithm 4 presents a solution to define speed of sound in the shell, using a look-up table of pre-calculated values based on the sphere's radius.

First an array of c_{pl} from $2500m/s$ to $10000m/s$ is defined, C_{pl_set}. The array is used to calculate a set of Δt_{shell} values, with a given radius a and speed of sound in the outer medium c_1, Equation 5.8. The value from the C_{pl_set} which provides minimum error with the measured Δt_{shell} is defined as c_{pl}.

To eliminate some fault results caused by a mistake in the plate wave component allocation, c_{pl} values outside the expected range ,c_{range}, are excluded. The expected range for the plate speed depends on the materials from the look-up table. The plate speed in the look-up table is in a range $(3879-5438)m/s$. Therefore in this study the expected range is defined as $(3000-6500)m/s$. In the "outside the range" case, $c_{pl} = NaN$ and material identification is impossible. In this case, large errors in the plate speed evaluation are detected automatically.

Algorithm 4 Speed of sound in the material

Input a, Δt_{shell}, c_1, c_{range}

Output c_{pl}

define $C_{pl_{set}} = 2500 : 10000$
for $n = 1 : length(C_{pl\_set})$ **do**
    $\theta_c = arcsin(\frac{c_1}{C_{pl\_set}(n)})$
    $T_{set}(n) = 2a(\frac{1-cos\theta_c}{c_1} + \frac{\pi-\theta_c}{C_{pl\_set}(n)})$
end
$pos = min(T_{set} - \Delta t_{shell})$
$c_{pl} = C_{pl\_set}(pos)$
if $c_{pl} <= c_{range}(1) or c_{pl} >= c_{range}(2)$ **then**
    $c_{pl} = NaN$
end

The shell material is defined based on the calculated plate speed in the shell, c_{pl} comparing the value with a look-up table of expected metals. Material with a nearest c_{pl} value is determined as the shell material.

5.3 Filler material evaluation

Filler material of a 2-layer sphere can be defined using the timing of the wave going through the sphere and reflected from the back wall. The time delay between specular reflection and

a reflection from the back wall of the sphere provides information about the speed of sound in the filler material. The filler liquid is defined based on the calculated speed of sound. The approach is similar to shell material identification and can be described by the scheme on the Figure 5.10.

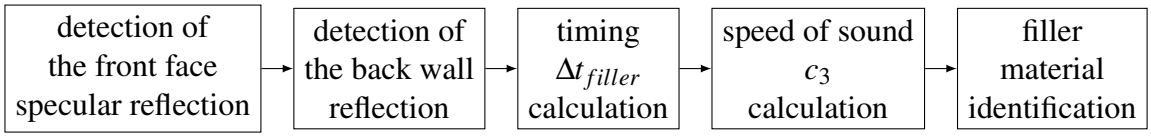


Figure 5.10. Scheme of the filler material identification

The filler identification approach is different from the shell material identification in the second and the fourth blocks of the scheme, where the back wall reflection is detected and the speed of sound in the liquid is calculated.

5.3.1 Path of the wave reflected from the back wall

The sound wave reflected from the back wall is propagating in two mediums: shell material and filler material. The path of the sound is presented on the Figure 5.11. There are other paths which the sound wave follows reflecting from the back wall, but these paths follow different direction and don't contribute into the component.

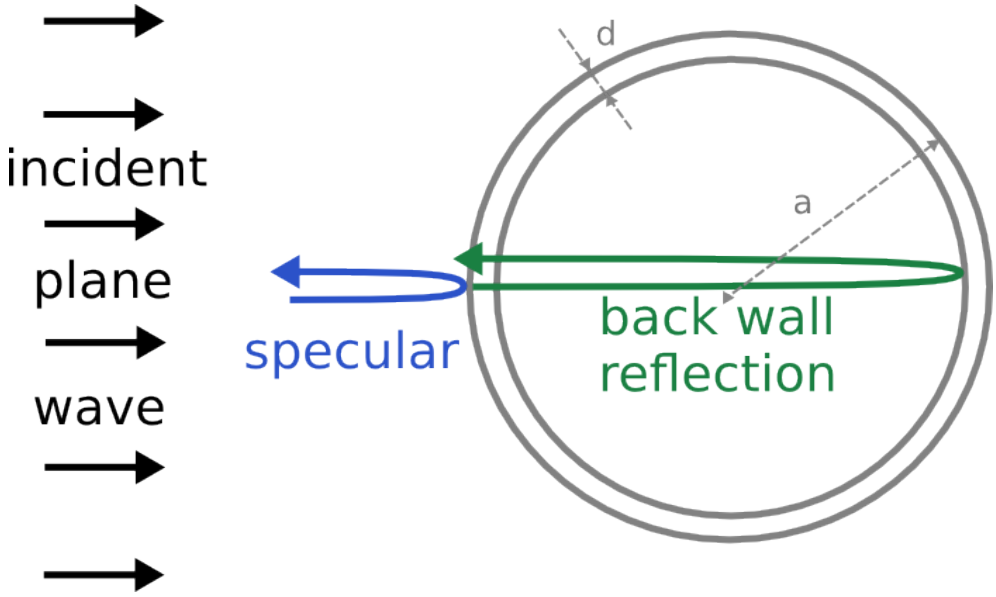


Figure 5.11. Trajectory of the wave reflecting from the back wall

The path of the wave reflected from the back wall can be divided into three components: (1) path through the shell, distance d, (2) path through the filler to the back wall with return path, distance $4(a-d)$, and (3) path back through the shell material, distance d.

The timing of the wave path is described by the Equation 5.10, where the first component is related to the propagation through the shell's material and the second component is the propagation through the filler.

$$t_{filler} = 2(\frac{d}{c_2} + \frac{2(a-d)}{c_3}), \tag{5.10}$$

where a is a sphere radius, d is a thickness of the sphere, c_2 and c_3 speed of sound in the shell material and filler material.

The timing for different sphere sizes and filler materials is presented in Table 5.3. It is calculated based on the Equation 5.10.

Table 5.3. Numerically calculated timing of the wave reflected from the back wall (different shell materials).

Materials	Δt_{shell} (μs)						
sphere radius	$0.05m$	$0.075m$	$0.1m$	$0.125m$	$0.15m$	$0.175m$	$0.2m$
aluminium sphere							
water	133	199	265	332	398	465	531
salt water	128	193	257	321	385	449	514
oil	137	206	275	343	412	481	550
kerosene	149	223	298	372	446	521	595
brass sphere							
water	133	199	266	332	399	465	532
salt water	129	193	257	321	386	450	514
oil	138	206	275	344	413	481	550
kerosene	149	223	298	372	447	521	596

The values are calculated for a sphere with a wall thickness 1% and two shell materials. Speed of sound in the liquids is provided in the Table 3.3, Chapter 3.

Difference between timing of the fillers is evident even for the pair "water - salt water", where the difference between their sound speeds is $50m/s$. The disparity increases, with the radius value, because of the increasing distance to propagate in the filler. That means a sphere with a bigger radius would be an easier target for the filler identification.

The type of the shell material doesn't really influence the timing for the thin shell with thickness less then $1-2\%$ of their radius. The path of the wave through the wall material is very short and with the speed of sounds a few times higher than the speed of sound in the filler liquid, it bring a very small difference. Table 5.4 illustrates changes in the timing while vary the shell thickness. The values are numerically calculated based on the Equation 5.10. The timing is slowly decreasing while thickness grows.

The wall thickness is an unknown parameter of the object. Based on the observation, that the propagation through the shell wall doesn't contribute much into the timing, the thickness of the wall can be a fixed value for the calculations.

Based on the Equation 5.10, the speed on sound in the filler can be defined from the timing Δt_{filler}, sphere radius a and speed of sound in the shell material c_2.

5.3.2 Timing of the wave reflected from the back wall

The time duration of the wave reflected from the back wall is a delay between specular reflection, $t_{specular}$, and the reflection from the back of the sphere's wall ,t_{wall}:

Table 5.4. Numerically calculated timing of the wave reflected from the back wall (different filler materials)

sphere radius	Δt_{shell} (μs)									
shell thickness	0.0%	1.0%	1.5%	2.0%	2.5%	3.0%	3.5%	4.0%	4.5%	5.0%
water filler										
0.05m	135	134	133	133	132	132	131	130	130	129
0.10m	270	268	267	265	264	263	262	261	260	258
0.15m	405	402	400	398	396	395	393	391	389	388
0.20m	541	536	533	531	529	526	524	521	519	517
salt water filler										
0.05m	131	130	129	128	128	127	127	126	126	125
0.10m	261	259	258	257	256	255	253	252	251	250
0.15m	392	389	387	385	384	382	380	378	377	375
0.20m	523	518	516	514	511	509	507	504	502	500

$$\Delta t_{filler} = t_{wall} - t_{specular} \tag{5.11}$$

The timing of the specular reflection is found using Algorithm 2, described in the Section 5.2. Timing of the wave which propagates through the sphere and reflects from the back wall of the shell is described by Algorithm 5. The approach allows to allocate peak related to the sound reflected from the back wall and calculate timing of the process t_{wall}.

The timing t_{wall} is calculated from the radius a, speed of sound in the outer medium c_1 and plate wave of the shell c_{pl}, reflected echo $s_{echo}(t)$ and coefficient k_{mpp}. The coefficient tunes the peak detector sensitivity to the prominence of the peaks in the matched filtering signal. The coefficient k_{mpp} defines a ratio of a minimum peak prominence (MPP) in relation to the matched filtering signals values. The matched filtering output $s_{mf}(t)$ is searched for peaks in a range from t_{2a} to t_{5a}. The time values represent a duration which the sound wave needs to propagate in the outer medium for $2a$ and $5a$ distances. The range is defined empirically.

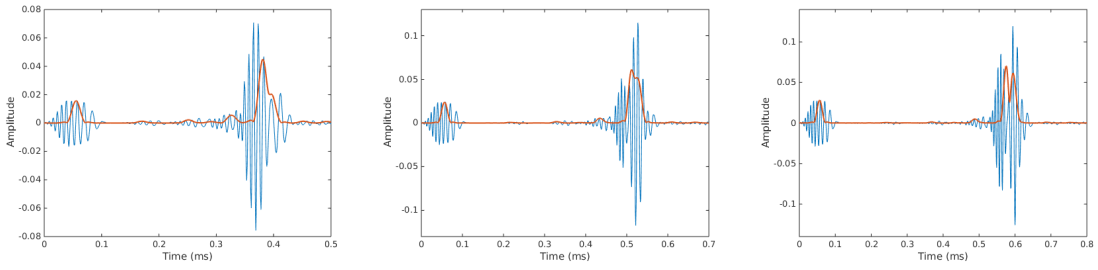
This approach provides adaptive tuning of the peak detector based on the matched filtering output. When the number of the detected peaks is two or less, MMP values is reduced until more peaks are found or MMP becomes 1% from the original value. In case of $MPP >= k_{MPP} * max(s_{mf}(t)) * 0.01$ the peak detector is stopped and t_{wall} value cannot be detected. When there are more than two peaks are detected, the set of peaks are reviewed to merge peaks which are too close to each other. This merge of peaks helps to avoid a confusion when a peak, related to the notches in frequency domain, is divided into two peaks. These double peaks appear while the radius of the sphere increases. The nature of the peak division is not clear, but can be related to the appearance of some new reflection components. Figure 5.12 illustrates appearance of the doubled peak. The data is analytically calculated for the stainless steel sphere filled with water. These doubled peaks can lead to a faulty peak detection.

Algorithm 5 Timing of the wave reflected from the back of the sphere

Input a, c_1, c_2, $s_{echo}(t)$, $s_i(t)$, k_{mpp}

Output t_{wall}

matched filtering $s_{mf}(t)$
calculate $t_{2a} = \frac{2a}{c_1}$ and $t_{5a} = \frac{5a}{c_1}$
calculate $MPP = k_{MPP} * max(s_{mf}(t))$
find peaks P_{loc} and P_{val} in $s_{mf}(t_{2a} : t_{5a})$
while $length(P_{loc}) <= 2$ & $MPP >= k_{MPP} * max(s_{mf}(t)) * 0.01$ **do**
    $MPP = MPP - 0.01MPP$
    find peaks P_{loc} and P_{val} in $s_{mf}(t_{2a} : t_{5a})$
    if $MPP < k_{MPP} * max(s_{mf}(t)) * 0.01$ **then**
        Position of the largest peak $N_{wall} = 0$
        break
    end
end
if $length(P_{loc}) > 2$ **then**
    if $\Delta(P_{loc}^{(N+1)} - P_{loc}^{(N)}) <= \frac{\pi a f_s}{2c_p l}$ **then**
        $P_{loc}^{merge} = (P_{loc}^{(N+1)} + P_{loc}^{(N)})/2$
        $P_{val}^{merge} = (P_{val}^{(N+1)} + P_{val}^{(N)})/2$
    end
    find the largest peak: $N_{max} = max(P_{val})$
    $N_{wall} = N_{max} - 1$
    calculate $t_{wall} = P_{loc}(N_{wall}) + t_{2a}$
else
    print " t_{wall} is not found"
end



$a = 0.10m$ $a = 0.14m$ $a = 0.16m$

Figure 5.12. Part of the echo reflected from a stainless steel sphere with different radius values: : recorded echo in time domain (blue) and matched filter output (red)

The figure presents a zoom into the first part of the echo, while there are more components following. The scale of the signal provides a detailed view of the peaks. In order to study the problem, data was simulated for an initial pulse with a half a bandwidth of the original pulse, Figure 5.13. This narrow bandwidth (generated in $(60-125)kHz$ range) leads to wider peaks, which reduce the effect and can improve the material identification.

The improvement is visible for this doubling peak problem, but the solution cannot be used

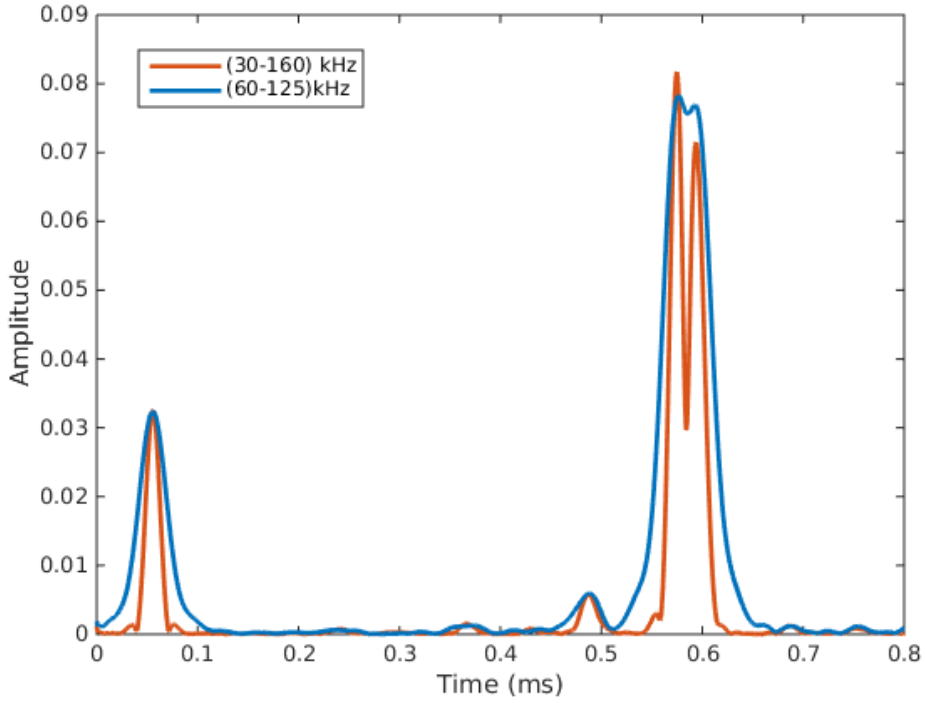


Figure 5.13. Changes in matched filtering output caused by reduction of the frequency range of the initial pulse: $(30-160)kHz$ - red and $(60-125)kHz$ - blue

on all the examples, as it influences the pulse detector and demotes the filler and shell material identifier results for other cases. This solution can be used if an adaptive approach is implemented to the task, when the pulse is changed based on the feedback from the identifier. In this case, if the doubled pulse is detected, the frequency range of the initial pulse will be reduced. It will lead to a smother matched filter signal with wider peaks, and as a result to reduction of the doubled peak problem. The topic is discussed further in Chapter 6, as part of the future work suggestion. In this work another solution was applied.

It is observed, that these doubled peaks are located much closer to each other, than others. The approach reviews the set of detected peaks and if the distance between peaks is small enough - merges them and consider the peaks as one. In this case, position and value of the merged peak is a mean value of these neighbouring peaks. The minimum distance is related to the sphere radius, a and c_{pl} of the sphere's material, as the peaks are related to the Lamb wave in the shell. The ratio of the parameters is adjusted using an experimentally based approach. The distance between peaks is limited by $\frac{\pi a f_s}{2c_{pl}}$.

When the peaks are reviewed and merged, the next step is to allocate a peak related to the distance between notches in frequency domain. The straight forward approach for the peak detection is to evaluate spacing between the notches, Δf, and relate it to the timing of the peak. It requires automatic evaluation of the spacing Δf, which is not a trivial task and for many example the peak detection is failing. This approach hasn't provided stable timing. Based on the observation of the data, the largest peak in the detection range (from t_{2a} to t_{5a}) is the peak related to the distance between notches in frequency domain. Therefore the

peak position of the largest peak, N_{max} is found based on the peak values array P_{val}. This approach provides reliable results in most of the cases.

One peak before the largest peak is related to the wave reflecting from the back wall of the shell. The position of the peak is $N_{wall} = N_{max} - 1$. The timing of the peak can be calculated as $t_{wall} = P_{pos}(N_{wall}) + t_{\frac{a}{2}}$.

Timing of the wave going though the sphere Δt_{filler} is calculated by Equation 5.11 based on the evaluated $t_{specular}$ and t_{wall}.

5.3.3 Speed of sound in the filler

The speed of sound in the filler and timing Δt_{filler} are related through Equation 5.10. This equation can be reformulated to express the speed of sound in the filler c_3:

$$c_3 = \frac{4c_2(a-d)}{\Delta t_{filler} c_2 - 2d} \tag{5.12}$$

The thickness of the shell, d, is the unknown parameter, as only radius of the sphere , a, is provided for the calculation. The shell thickness value is fixed for all the spheres and it is assumed to be 1% of the sphere radius, a. The value of the speed of sound in the shell material is taken based on the shell material identification in the Section 5.2. If the shell material is not defined, then it is assumed to be aluminium. These assumptions reduce the accuracy of the approach, but allows the calculation without any prior knowledge about the sphere's thickness.

The calculated speed c_3 is compared to the expected range and the values outside the range are excluded. The expected range depends on the look-up table. Speed of the filler material in the look-up table is from $1380 m/s$ (kerosene) to $1530 m/s$ (salt water), therefore the expected range for the experiment is chosen as $(1200 - 1600) m/s$. For the calculated speed outside the range, the value is reassign $c_3 = NaN$ and material identification is impossible.

The filler material is identified based on a look-up table. It is the one with the closest speed of sound value to the calculated result.

5.4 Results

The approaches for the material identification of the sphere's shell and filler is evaluated based on the results for the synthetic data, Section 5.4.1, and experimental data, Section 5.4.2. The results are evaluated in the order of the material identification approach. The results are presented firstly for the timing evaluation, where mainly peak-detector performance determines the accuracy of the method. Results for the calculation of the sound speed are presented next, they are similar to the time evaluation results, as they also depends on the peak detector, but influenced by assumptions made for the speed calculation as well. The material identification approach itself is evaluated at last. The results are pre-

sented differently for the synthetic and experimental data, due to the difference in variety of objects and materials.

5.4.1 Synthetic data

The synthetic data is calculated based on the approach described in Chapter 3, where analytical calculation of a wideband echo reflected from a spherical object immersed in water is presented. The reflected echoes are calculated for a number of spherical objects with variety of radii (from $0.05m$ to $0.15m$), thickness of the shell (from 1% to 2% from the radius value), shell material (aluminium, stainless steel, brass, copper) and filler material (water, salt water, oil and kerosene). In total there are 528 echo samples, each material is presented by 132 samples. Each echo is processed using the proposed approach to evaluate shell and filler material.

The results of the timing evaluation are presented in the form of graphs with a ground truth lines for each material. The ground truth is numerically calculated based on the Equations 5.13 and 5.14. The timing results are shown for a set of spheres' radius. As the thickness of the sphere varies as well, each radius value is presented by three points referring to the thickness of 1%, 1.5% and 2%.

Speed of sound results are shown by graphs and a table. The table combines results of all the objects based on their material and presents the average of the evaluated speed, which is the same for all the objects sharing the same material. The table representation of the results is not possible for the timing evaluation, where timing varies with the sphere's size. The table highlights the mean value and standard deviation (std) of the speed for each material with a percent error regarding the ground truth (GT) value. The ground truth for the speed of sound is taken from the Tables 5.2 and 3.3.

The material identification is evaluated with a matching matrix and a success rate over all the results. As the materials are identified from a look-up list, where only expected materials with a known speed of sound can be identified, the matching matrix can be used. The success rate is calculated for all the 528 samples and it presents a percent of true positives over all the set of samples.

Shell material

Results for the plate wave timing evaluation are presented in Figures 5.14 and 5.15. The ground truth (GT) is numerically calculated based on the known parameters of the spheres, Equation 5.13.

$$\Delta t_{shell} = 2a\left(\frac{1-cos\theta_c}{c_1} + \frac{\pi-\theta_c}{c_{pl}}\right), \tag{5.13}$$

The timing Δt_{shell} is increasing with the radius. The results are presented separately for four filler cases: water, salt water, oil and kerosene, Figure 5.14. The ground truth on the graph is presented by line and the described approach is presented by diamonds. Four

different shell materials are showed on each graph: aluminium, stainless steel, copper and brass. The results are presented for a set of radii (x-axis).

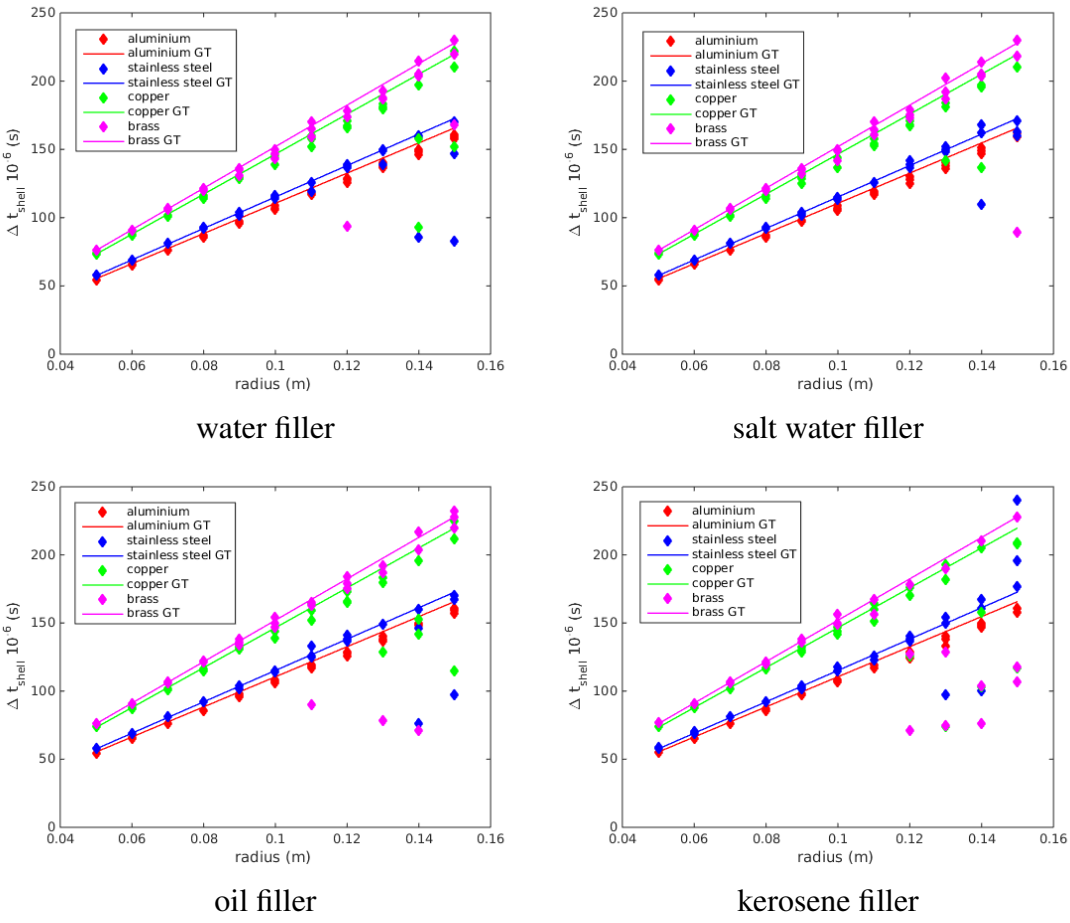


Figure 5.14. Timing of the plate wave Δt_{shell} for different shell materials (aluminium, stainless steel, copper, brass) illustrated for four different fillers (water, salt water, oil and kerosene). Results are presented for the synthetic data, GT is calculated numerically

The approach results (diamonds) follows the GT lines and difference between shell materials for the same radius value is noticeable. Even though evaluated timing doesn't fit the GT line precisely, it is close for most of the samples. Timing of the smaller radius up to $0.1m$ is evaluated more precisely, than the rest. For the higher values of the radius, there are number of outlying results, which presents a clear error in the peak detection. It can be related to the longer path of the plate wave for the bigger radius, which brings higher attenuation of the plate wave component and makes the component detection difficult.

Figure 5.15 presents results of plate wave timing evaluation for different fillers and the same shell material. These results show how much the shell material evaluation is influenced by the filler. The GT for the filler materials is located at the same line, as there is no expected influence of the filler material into the timing of the plate wave. The results are presented for four shell materials separately.

The plots don't show any noticeable trend related to the filler material. The results are following the GT line, but do not fit the ground truth completely due mainly to the errors in peak detection. The timing of the aluminium shells is evaluated with the highest accuracy

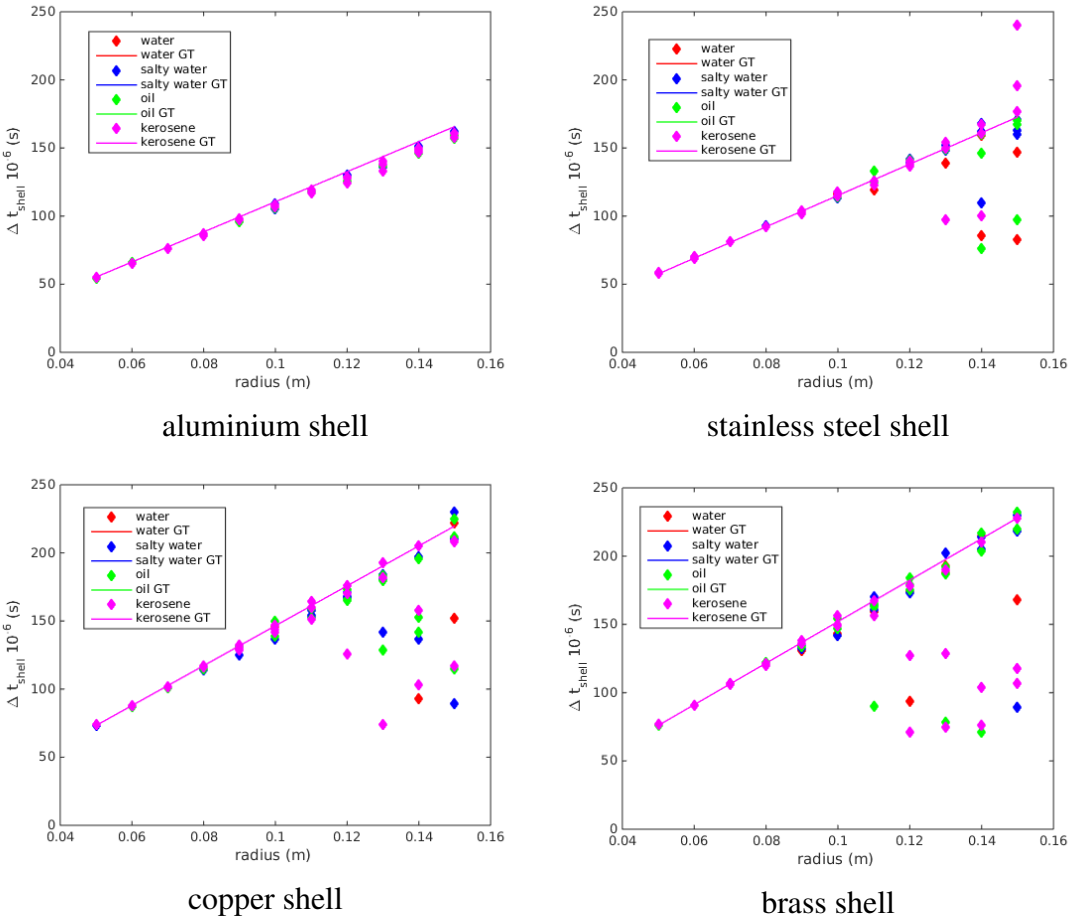


Figure 5.15. Timing of the plate wave Δt_{shell} for different filler materials (water, salt water, oil, kerosene) illustrated for four different shell materials (aluminium, stainless steel, copper, brass). Results are presented for the synthetic data, GT is calculated numerically

and without large errors, which can be related to the properties of the material and the way the plate wave is propagating in the shell. Based on the results, it can be concluded that the filler material has no visible influence on the Δt_{shell} timing evaluation by the presented method.

The speed of sound of the plate wave, c_{pl} is defined based on the evaluated timing. The ground truth is the plate wave value of the material, which is described in Table 5.2. Results for the speed of sound is presented in Figures 5.16 and 5.17, as well as in Table 5.5. The results on the graphs are presented for the different shell materials and different filler materials separately the same way as it was done for the timing result evaluation. The ground truth lines are horizontal, as the sound speed doesn't vary with the radius and thickness. In Figures 5.16, comparing the speed of sound c_{pl} for the same radius, it can be seen that the difference between the shell materials is visible, even though the diamonds of the method results are not always located on the GT lines. The c_{pl} results are similar for all four cases with different fillers: the plate speed is evaluated more accurately for the spheres with smaller radius. The results are calculated from the timing and the reasons for the performance are described in the timing results.

Figure 5.17 shows the results for the same shell material filled with different fillers. One

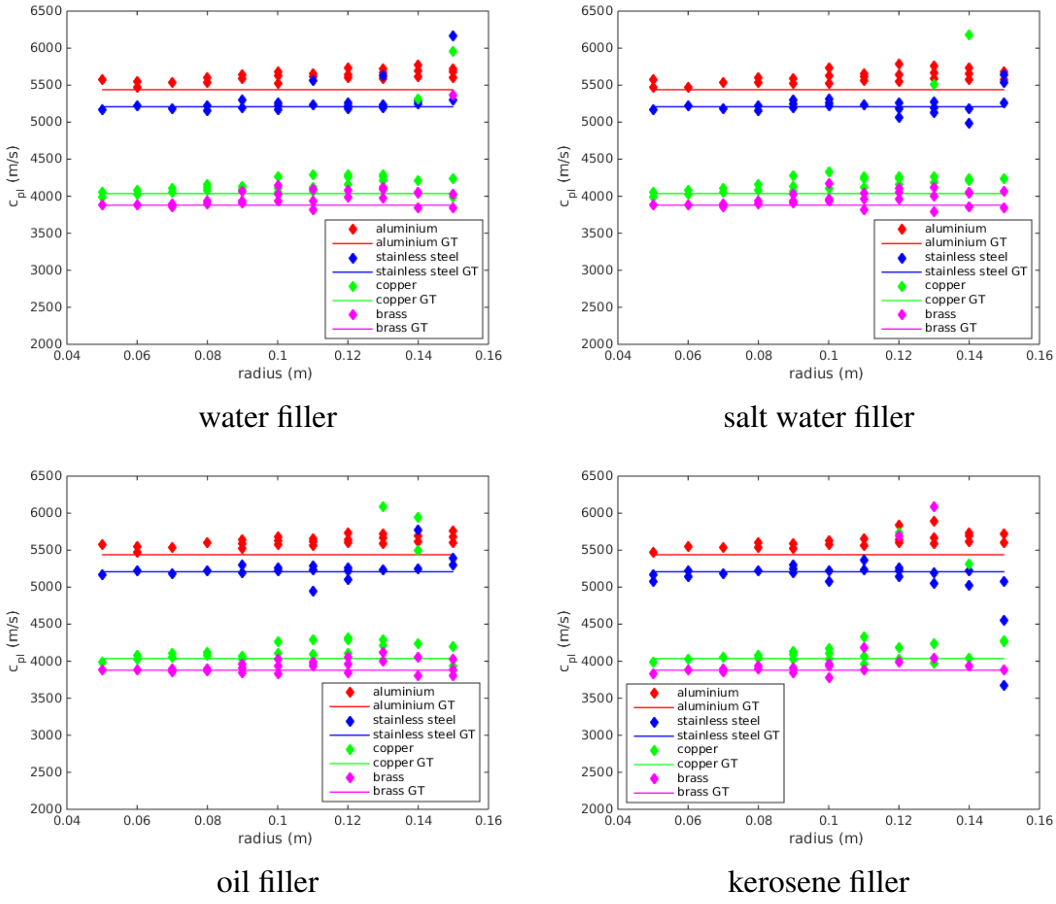


Figure 5.16. Plate speed of the shell material c_{pl} for different shell materials (aluminium, stainless steel, copper and brass) illustrated for four different fillers (water, salt water, oil, kerosene). Results are presented for the synthetic data, GT is calculated numerically

GT line presents the results for all four filler materials, because there is no influence of the filler material for the GT. And as it would be expected based on the timing evaluation there is no influence for the approach as well. Results of aluminium sphere doesn't show any big errors in the plate speed evaluation, unlike stainless steel, copper and brass. The reasons are described in the results of timing evaluation.

This performance of the speed evaluation is expected as the results directly depends on the timing evaluation. It can be seen that there are cases with some visible large errors, where the calculated speed is close to the wrong material. This will bring to the wrong identification on the next step. Nevertheless majority of the results fit the GT of its own material. To evaluate the error of the speed evaluation and provide numerical evaluation of the presented graphs, the average results are presented in Table 5.5.

Table 5.5. Synthetic data results: plate speed in the shell material

spheres		speed of sound c_{pl}		
N	shell material	calculated (m/s)	experiment (m/s)	error (%)
1	aluminium	5438	5603 ± 82	3.0
2	stainless steel	5208	4931 ± 1187	5.3
3	copper	4035	4038 ± 967	0.1
4	brass	3879	3646 ± 1141	6

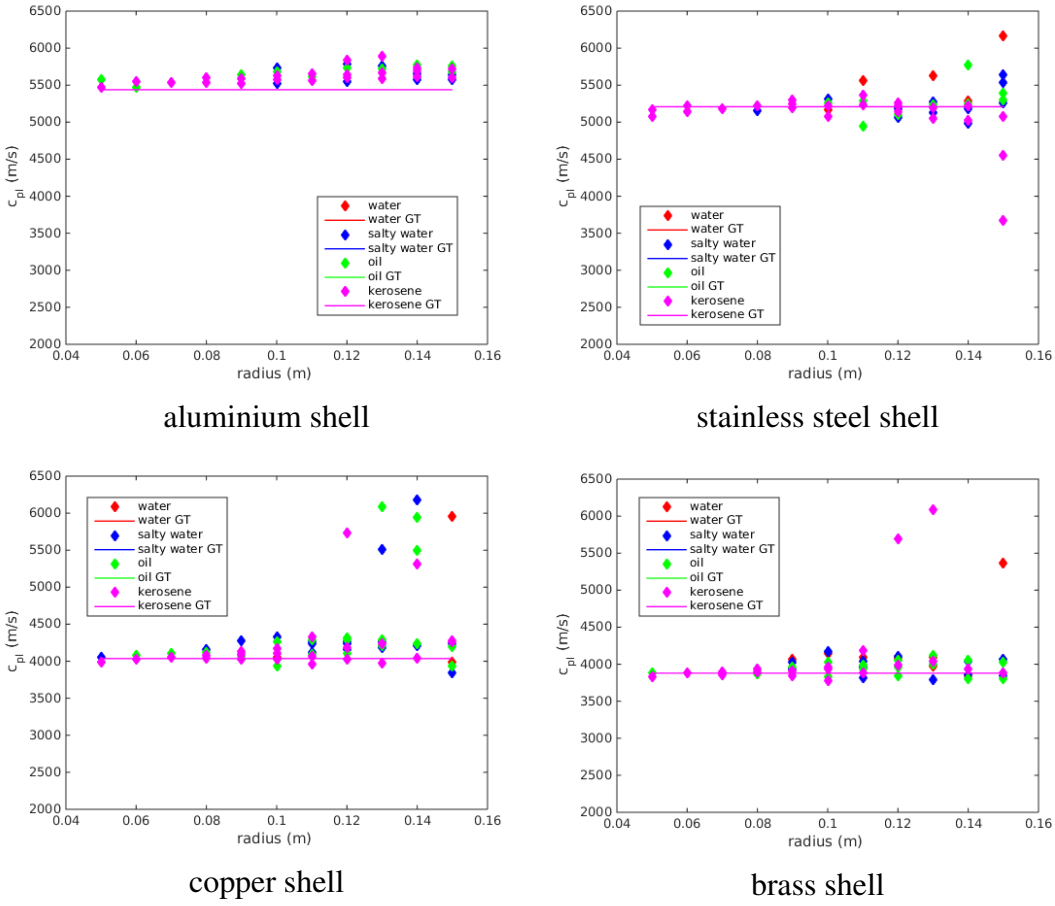


Figure 5.17. Plate speed of the shell material c_{pl} for different filler materials (water, salt water, oil, kerosene) illustrated for four different shell materials (aluminium, stainless steel, copper and brass). Results are presented for the synthetic data, GT is calculated numerically

The error doesn't exceed 6% after elimination of the results outside the expected plate speed range $(3000-6500)m/s$. The aluminium shell material performs with the smallest std, which shows that the method provides stable results for the set of radius and thickness. The copper shell results have a higher accuracy, based on the average c_{pl} with a high std values.

All the shell material identification results are grouped together for the evaluation. The matching matrix is presented in Table 5.6. There are 132 samples of each material.

Table 5.6. Matching matrix of shell material identification with 132 samples of each material

shell material	aluminium	stainless steel	copper	brass
aluminium	132	8	8	3
stainless steel	0	115	1	0
copper	0	1	114	38
brass	0	1	3	80

Aluminium is identified with 100% true positives. The material identification approach shows the lowest performance for the brass material, as a lot of samples were wrongly detected as copper. The speed of sound of copper and brass is close to each other and it gives an error in identification even for a small error in the calculated c_{pl}.

The performance of method for the shell material identification is also evaluated by the success rate. The success rate is calculated for the generated dataset of 528 spheres, which includes different shell materials (aluminium, stainless steel, copper and brass), different filler materials (water, salt water, oil and kerosene), sphere's radius in a range $(0.05 - 0.15)m$ and shell thickness $(1-2)\%$. The shell material is identified with a success rate 83.5%.

Filler material

The timing of the filler material identification is presented in Figures 5.18 and 5.19. The ground truth is calculated based on the sphere's parameters, Equation 5.14.

$$t_{filler} = 2(\frac{d}{c_2} + \frac{2(a-d)}{c_3}), \tag{5.14}$$

Timing for different filler materials is presented in Figure 5.18. The figure shows graphs for the four shell materials: aluminium, stainless steel, copper and brass. The GT is presented by lines and the approach results by diamonds. Timing value is growing proportional to the radius. It can be seen, that apart from a few points the evaluated timing Δt_{filler} fits well the GT lines. The difference in the timing for the same radius and different filler materials is clearly visible. To evaluate how shell material influences the results for the Δt_{filler}, Figure 5.19 shows graphs where timings for different shell material and the same filler are presented. Only one GT line is visible on the graphs, as it is expected that the shell material will not influence the timing and all the lines are overlapped. The results for the variation in the shell materials align well with the GT apart from a small number of points. These outlying points relate to the wrong peak detection, which mostly occurs, because of the peak division described above in the Section 5.3. It can be concluded, that the approach for the filler material evaluation is not influenced by the shell material.

Results of the speed of sound in the filler are presented in Figures 5.20 and 5.21, and numerically in Table 5.7. The ground truth (GT) for the speed of sound is presented in Table 3.3, Chapter 3. In the approach, the speed of sound is calculated based on the timing of the wave, therefore the same behaviour of the results is expected. Figure 5.20 shows graphs with speed of sound c_3 for different filler material and the same shell material. The timing values for the same radius and different filler material show clear separation of the filler materials. There are a number of outlying points, which presents detection of a wrong peak, which are related to the secondary peak division. Graphs with the same filler materials, but different shell materials, Figure 5.21, demonstrate no visible influence of the shell material into the filler identification. The GT line and approach results are overlapping each other for different shell materials. It can be seen that spheres with a small radius have a lower accuracy, which can be related to a shorter path of the wave reflected from the back wall. The shorter path makes the approach more sensitive to the timing error. This timing error becomes visible after calculation of the c_3. Also the calculations are based on the assumption of the shell thickness 1%, which influences the accuracy of the speed

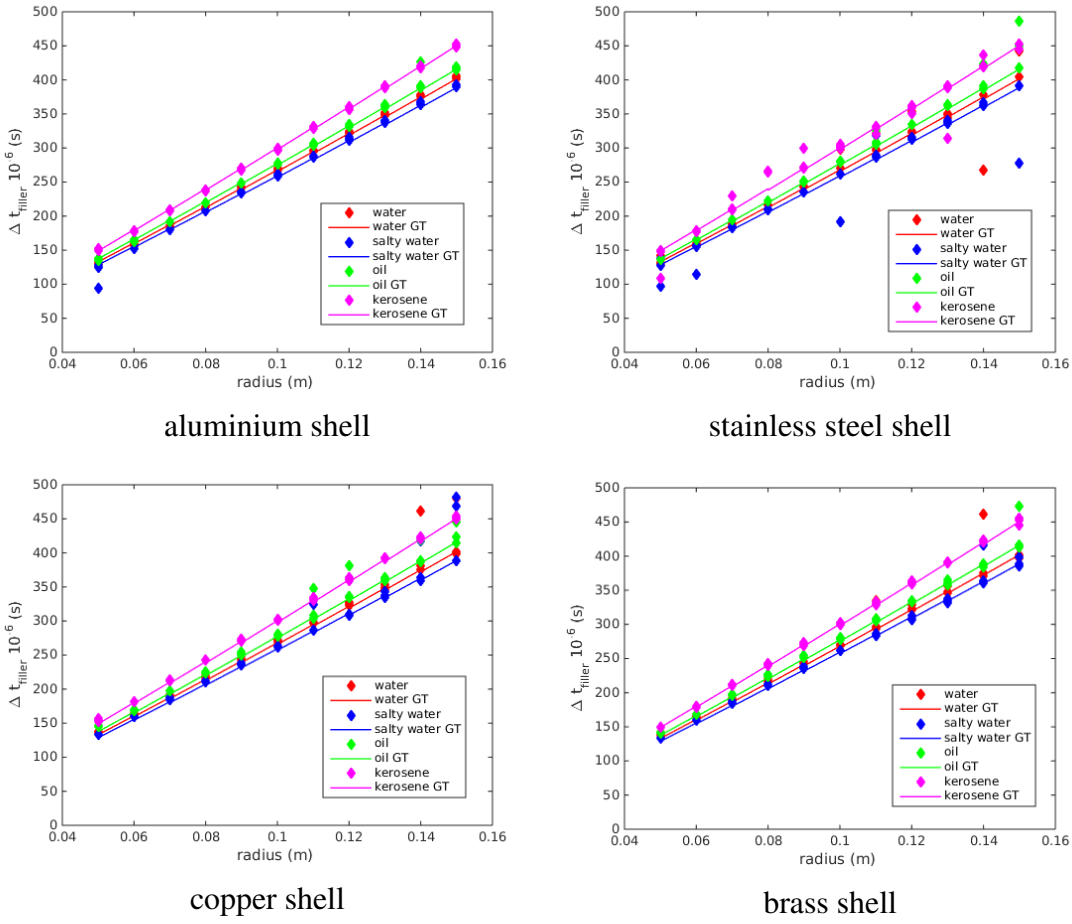


Figure 5.18. Timing of the wave reflected from the back wall t_{filler} for different filler materials (water, salt water, oil, kerosene) illustrated for four different shell materials (aluminium, stainless steel, copper, brass). Results are presented for the synthetic data, GT is calculated numerically

evaluation.

The results are also presented as the mean value and standard deviation (std) of all the measurements for the given material and compared to their original values. Table 5.7 shows the maximum error of 5.5% in speed evaluation, which is appeared for the salt water filler. This result is performed with a highest deviation as well. Speed in oil filler is evaluated with a lowest error and lowest deviation.

Table 5.7. Synthetic data results: speed of sound in the filler

spheres		speed of sound c_3		
N	shell material	GT (m/s)	experiment (m/s)	error (%)
1	water	1480	1449 ± 97	2.1
2	salt water	1530	1445 ± 221	5.5
3	oil	1430	1411 ± 39	1.3
4	kerosene	1320	1258 ± 136	4.7

The filler material of the sphere is identified based on the calculated speed of sound. The filler identification results are grouped together for the evaluation. The matching matrix is calculated with 132 samples for each material, Table 5.8. The kerosene filler is identified with the highest true positives. The reason is that salt water, water and oil have a similar

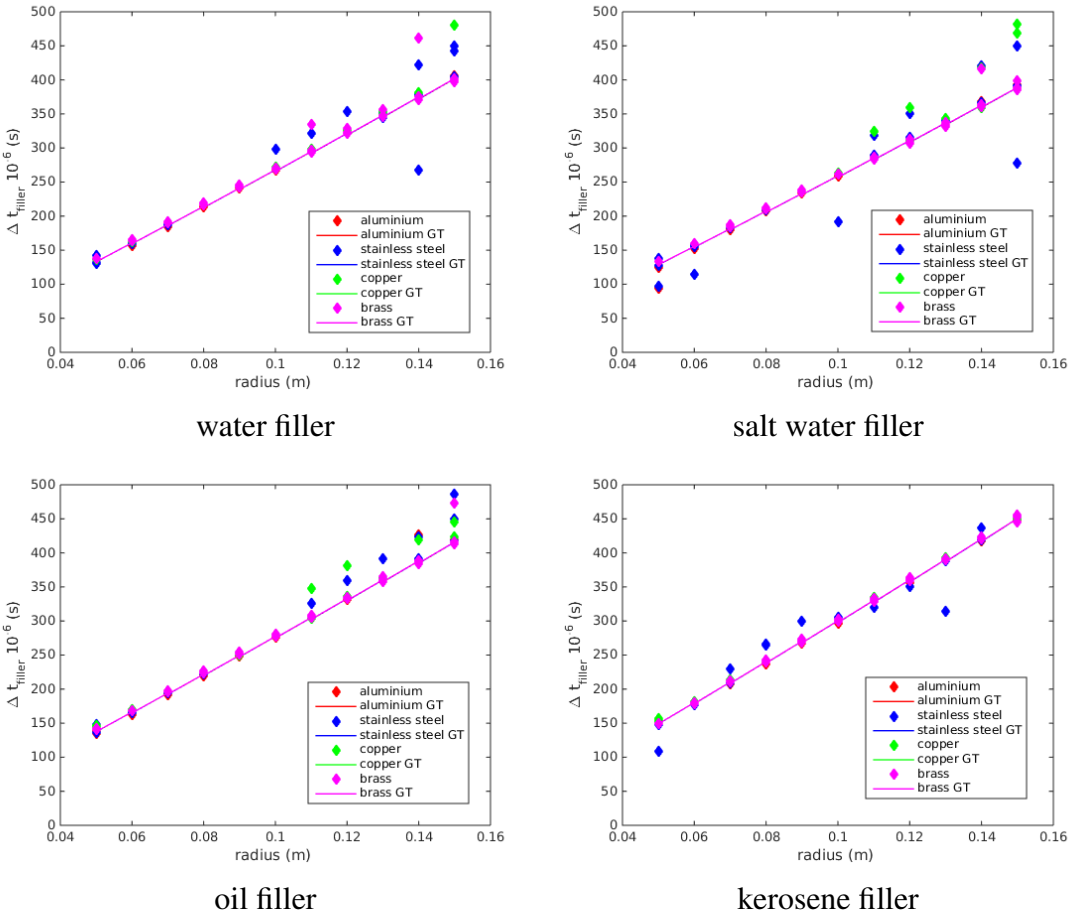


Figure 5.19. Timing of the wave reflected from the back wall t_{filler} for different shell materials (aluminium, stainless steel, copper, brass) illustrated for four different fillers (water, salt water, oil, kerosene). Results are presented for the synthetic data, GT is calculated numerically

sound speed, while kerosene difference by more then $100m/s$. It makes the kerosene identification less sensitive to the errors in c_3 value. For the oil filler, 16 samples were counted as kerosene, as well for the salt water 26 samples were counted as water.

Table 5.8. Matching matrix of the filler material identification with 132 samples of each material

shell material	water	salt water	oil	kerosene
water	95	26	5	0
salt water	5	90	0	0
oil	21	1	111	0
kerosene	10	10	16	126

The success rate of the filler identifier is calculated for the dataset of 528 spheres, which includes different filler materials (water, salt water, oil and kerosene), shell materials (aluminium, stainless steel, copper and brass), sphere's radius in a range $(0.05-0.15)m$ and shell thickness $(1-2)\%$. The filler material is identified with a success rate of 79.9%.

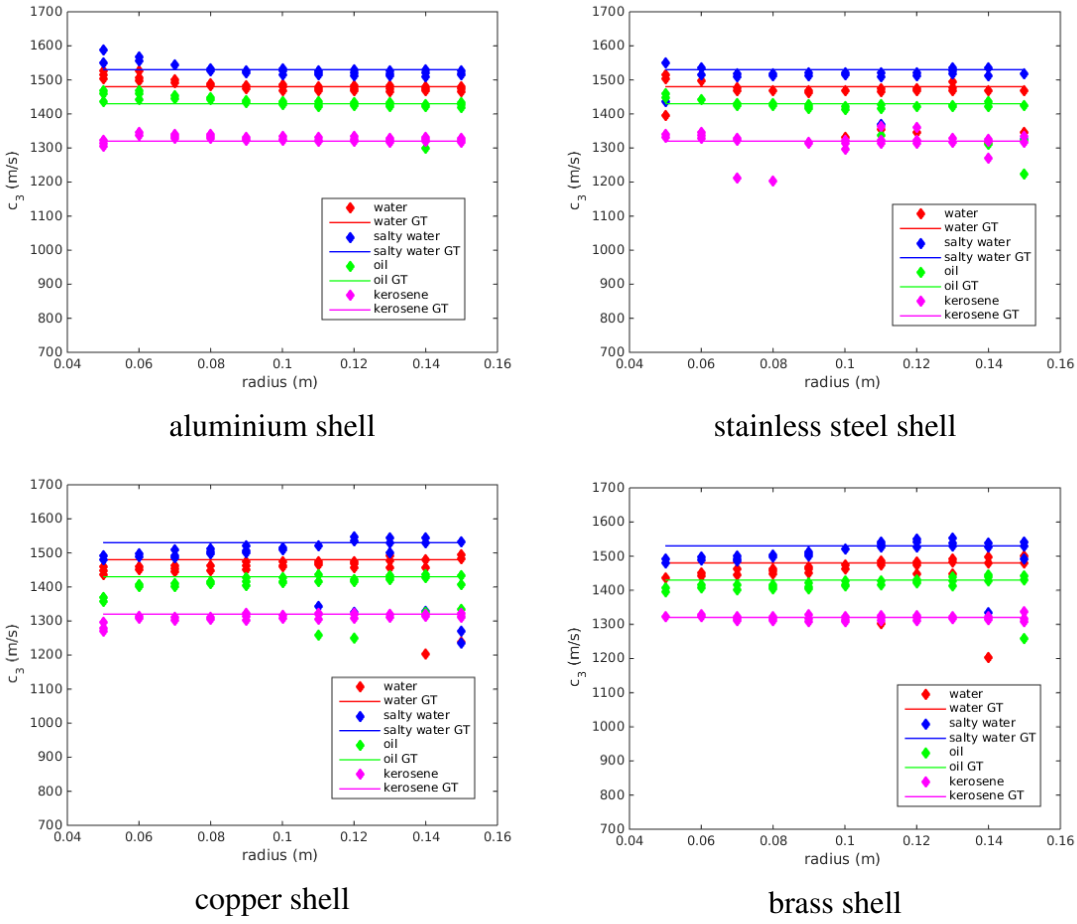


Figure 5.20. Speed of sound in the filler c_3 for different filler materials (water, salt water, oil, kerosene) illustrated for four different shell materials (aluminium, stainless steel, copper, brass). Results are presented for the synthetic data, GT is calculated numerically

5.4.2 Experimental data

The results for the timing and speed of sound evaluation are given in Tables 5.10 - 5.13, where the object is presented with its different fillers. The tables specify the mean value and std, based on the five independent echoes from the same target. Error represents a difference between GT and the result of the method in percentage. The final material identification is evaluated by the success rate.

All the experimental data is recorded in the test tank of the Ocean Systems Laboratory. The experimental set-up is described in the Section 4.1. The material identification approach is evaluated for four different spherical targets filled with three filler liquids.

Objects

The objects are three stainless steel spheres and one aluminium sphere filled with three different fillers, Figure 5.22. The aluminium sphere has rumples on its surface and 5 holes of different diameters. Spheres 2, 3 and 4 are made of a stainless steel material, each of them has two holes for the liquid refill. The described holes on the spheres are covered with a tape to prevent a leakage of the filler liquid in the test tank during the experiments.

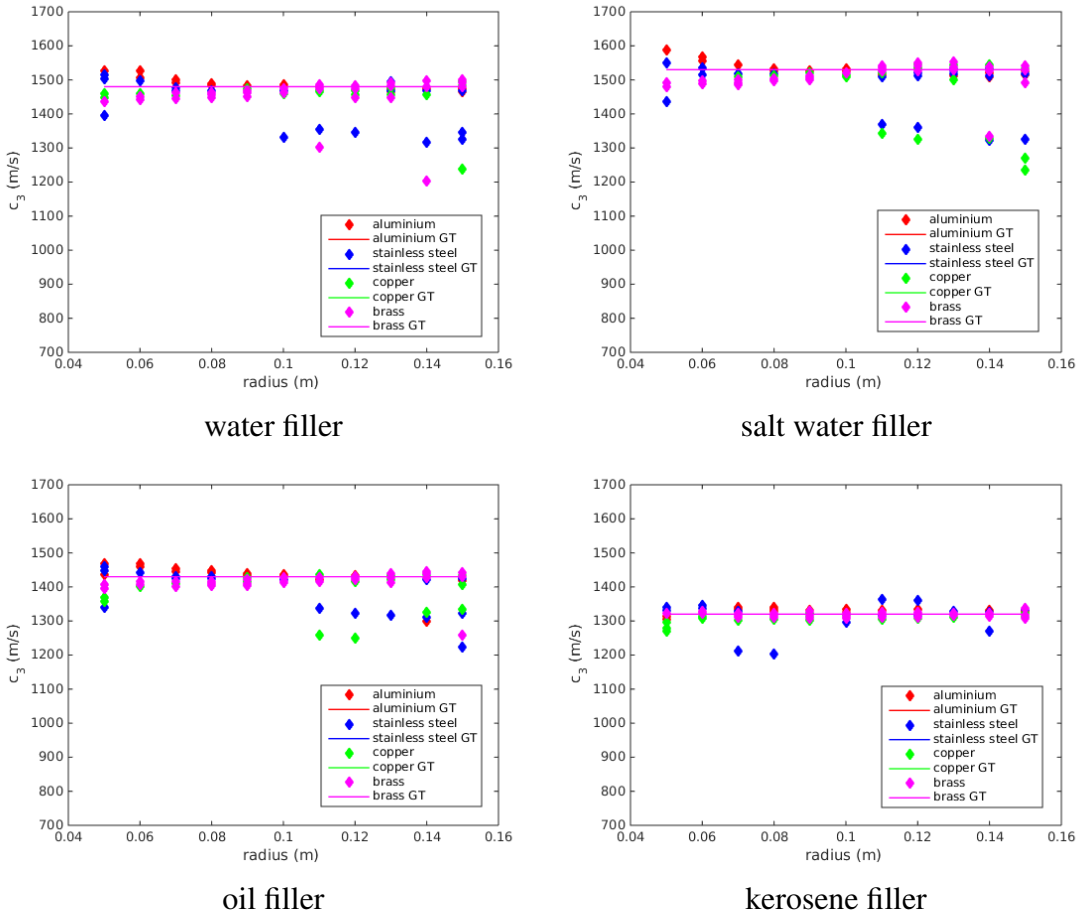


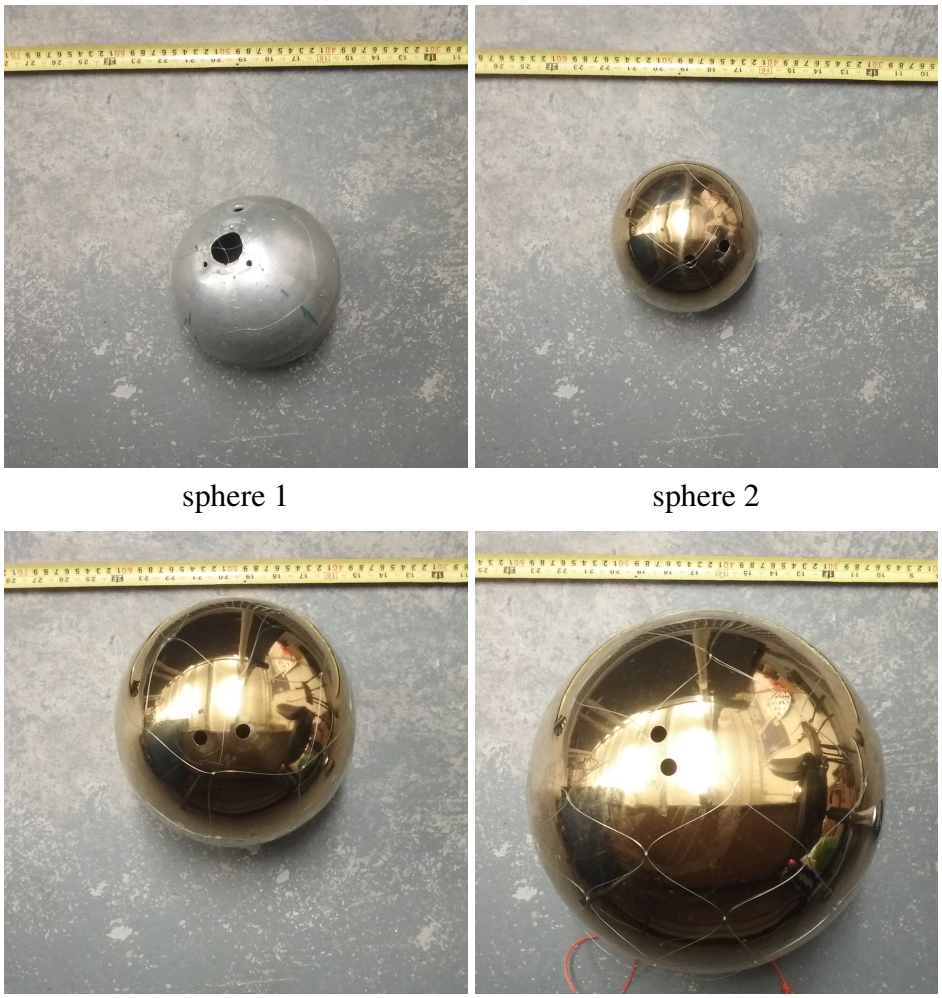
Figure 5.21. Speed of sound in the filler c_3 for different shell materials (aluminium, stainless steel, copper, brass) illustrated for four different fillers (water, salt water, oil, kerosene). Results are presented for the synthetic data, GT is calculated numerically

It brings some disturbances and slightly influences reflected echoes. The targets were filled with 3 different liquids: water, salt water and oil. Table 5.9 lists the spheres' size and materials. The spheres were filled with the liquids manually and the best attempt was made to leave no air sealing them.

Table 5.9. Description of the targets for the recognition

object	radius, m	shell thickness, m	shell material	filler
1	0.075	0.001	aluminium	fresh water, salt water
2	0.065	0.001	stainless steel	fresh water, salt water, oil
3	0.100	0.001	stainless steel	fresh water, salt water, oil
4	0.140	0.001	stainless steel	fresh water, salt water, oil

The objects were placed one by one in the test tank which was filled with fresh water. The echo from each object was recorded five times from different angles and distances to the objects. The results of the material identification approach are presented below separately for the shell and filler materials.



sphere 1 sphere 2

sphere 3 sphere 4

Figure 5.22. Objects for the experimental data recordings

Table 5.10. Experimental results: timing of the plate wave component

spheres			timing of the plate wave Δt_{shell}		
object N	shell material	filler	GT ($10^{-6} sec$)	experiment ($10^{-6} sec$)	error (%)
1	aluminium	water	83	89 ± 11	6.7
1	aluminium	salt water	83	83 ± 14	0.0
2	stainless steel	water	75	87 ± 10	16.0
2	stainless steel	salt water	75	77 ± 1	2.7
2	stainless steel	oil	75	81 ± 10	8.0
3	stainless steel	water	115	115 ± 2	0.0
3	stainless steel	salt water	115	115 ± 2	0.0
3	stainless steel	oil	115	115 ± 1	0.0
4	stainless steel	water	161	166 ± 12	3.1
4	stainless steel	salt water	161	183 ± 52	12
4	stainless steel	oil	161	157 ± 1	2.5

Shell material

The results of the plate wave timing evaluation is presented in the Table 5.10. The ground truth (GT) is calculated based on the Equation 5.8 from the sphere's size and material's characteristics. The experimental results are presented in a format of $(mean \pm std)$. The error is expressed in percent.

The approach provides low level of errors with the highest error presented for the case of the $0.065m$ radius stainless steel sphere filled with water, while $0.1m$ radius stainless steel sphere is evaluated with 0% error.

The speed of sound is calculated from the measured plate wave timing. The ground truth is a speed of sound in the materials taken from Table 5.2. The results are presented in Table 5.11.

Table 5.11. Experimental results: plate speed in the shell material

spheres			speed of sound c_{pl}		
object N	shell material	filler	GT (m/s)	experiment (m/s)	error (%)
1	aluminium	water	5438	5245 ± 620	3.5
1	aluminium	salt water	5438	5576 ± 1017	2.5
2	stainless steel	water	5208	4487 ± 587	13.8
2	stainless steel	salt water	5208	5056 ± 85	2.9
2	stainless steel	oil	5208	4843 ± 625	7.0
3	stainless steel	water	5208	5225 ± 78	0.3
3	stainless steel	salt water	5208	5235 ± 81	0.5
3	stainless steel	oil	5208	5206 ± 40	0.0
4	stainless steel	water	5208	5078 ± 365	2.5
4	stainless steel	salt water	5208	4832 ± 124	7.2
4	stainless steel	oil	5208	5384 ± 16	3.4

The plate speed evaluation results slightly differ from the timing evaluation due to the non-liner relation between the terms. The shell material is identified based on these calculated values. The shell's material is chosen from four possible variations: stainless steel, aluminium, brass and copper. The experimental data contains only two types from the list. The shell material identification approach is tested on the 55 test examples and it gives 67.3% success rate.

It is shown that the method provides a low percentage of errors, apart from a few cases highlighted above. The material identification depends on the look-up table of the possible shell's materials. Aluminium and stainless steel has a value of plate speed close to each other and even small error can influence the identification results. It decreases the success rate of the shell material identification.

Table 5.12. Experimental results: timing of the back wall reflection

spheres			timing of the back wall reflection Δt_{filler}		
object N	shell material	filler	GT $(10^{-6}sec)$	experiment $(10^{-6}sec)$	error (%)
1	aluminium	water	199	201 ± 2	1.5
1	aluminium	salt water	193	192 ± 4	0.5
2	stainless steel	water	173	177 ± 1	2.3
2	stainless steel	salt water	168	170 ± 6	1.2
2	stainless steel	oil	179	176 ± 1	1.7
3	stainless steel	water	268	271 ± 1	1.1
3	stainless steel	salt water	259	245 ± 16	5.4
3	stainless steel	oil	277	273 ± 1	1.4
4	stainless steel	water	376	378 ± 1	0.5
4	stainless steel	salt water	364	377 ± 1	3.6
4	stainless steel	oil	389	382 ± 1	1.8

Filler material

The results for the timing of the wave reflected from the back wall are presented in Table 5.12. The ground truth (GT) is calculated based on the Equation 5.10 from the sphere's size and material's characteristics. The approach provides stable result with a low percentage of error. The highest error occurs for the $0.1m$ radius stainless steel sphere filled with salt water, this result also has the highest standard deviation.

The speed of sound in the filler material values depends on the evaluated timing. The ground truth is a speed of sound in the filler materials from the look-up list, Table 3.3, Chapter 3. The difference between timing and speed results due to the number of assumptions in the speed calculation.

Table 5.13. Experimental results: speed of sound in the filler

spheres			speed of sound c_3		
object N	shell material	filler	GT(m/s)	experiment (m/s)	error (%)
1	aluminium	water	1480	1477 ± 12	0.2
1	aluminium	salt water	1530	1550 ± 28	1.3
2	stainless steel	water	1480	1456 ± 82	1.6
2	stainless steel	salt water	1530	1513 ± 54	1.1
2	stainless steel	oil	1430	1460 ± 3	2.1
3	stainless steel	water	1480	1461 ± 3	1.3
3	stainless steel	salt water	1530	1625 ± 100	6.2
3	stainless steel	oil	1430	1454 ± 7	1.7
4	stainless steel	water	1480	1466 ± 3	0.9
4	stainless steel	salt water	1530	1518 ± 4	0.8
4	stainless steel	oil	1430	1454 ± 2	1.7

The filler materials are identified from a list of four options: water, salt water, oil and kerosene. The material identification approach provides 80% success rate, which is calculated based on the 55 test examples.

5.5 Summary

In this chapter an approach for the shell and filler materials identification of a metal spherical shells filled with liquid was presented. This approach is based on the measurement of the echo component's timing. The developed method doesn't involve machine learning and requires no training data. The requirement for the prior knowledge is a look-up table with the sound speed in the materials (plate speed for the shell and longitudinal speed of sound for the liquid fillers), the sphere's radius and recorded wideband echo. The approach is limited by the 2% thickness assumption and tested for the diameter up to $0.3m$.

The components of the reflected echo were described in Section 5.1. The components, including specular reflection, plate wave and wave reflecting from the back wall of the shell are essential for the material identification. The trajectory of the components were calculated based on the sphere's size. The timing of the sound waves was extracted from the reflection based on the matched filtering technique and a peak detector. A basic adap-

tive control over the minimum peak prominence for the peak detector was integrated into the solution. The value is changed to detect a sufficient number of peaks for the timing evaluation. The method was presented separately for the shell material and filler material identification. Both methods are similar, their difference is based on the way how the peaks related to the components are selected and calculation of the sound speed.

The results for the material recognition methods were presented for synthetic and experimental data. The synthetic data includes objects with variety in radius (from $0.05m$ to $0.15m$), shell thickness (from 1% to 2% of the radius value), shell material (aluminium, stainless steel, brass, copper) and filler material (water, salt water, oil and kerosene). In total 528 variation of the spherical objects with 132 echo samples for each material are evaluated. Comparing results for different radii and materials for the shell material timing and plate speed, it can be concluded that the difference between shell materials is clearly visible and the filler material has no visible influence on the Δt_{shell} timing and speed c_{pl} evaluation. Reflection from the smaller radius spheres (up to $0.1m$) are evaluated more precisely, than for the larger spheres. For the higher values of the radius, there are a number of incorrect results, which presents clear an error in the peak detection. It can be related to the longer path of the plate wave for the bigger radius, which brings higher attenuation of the plate wave component and makes the component detection difficult. The plate wave of four metals is evaluated with an averaged error of 3.5%. The filler material identifier also presents a good fit into the ground truth for the timing and speed of sound values. The outlier points in the results are related to the wrong peak detection, which mostly occur, because of the peak division. There is a clear separation between the filler material and the shell material doesn't influence the filler identification results. The speed of sound in the filler is evaluated with an averaged error of 3.4% among all four fillers. The approach provides 83.5% success rate for the shell material and 79.9% success rate for the filler material identification with synthetic data.

Experimental data was recorded in the test tank and included four spheres. The spheres were made of stainless steel and aluminium. They were filled with three different liquids: water, salt water and oil. In total there were 11 objects for the shell and filler material evaluation. The echo for each object was recorded 5 times from different angles and distances to the objects. In this way, the methods applied for the experimental data was evaluated based on the 55 echo samples from 11 spheres. The plate wave was evaluated with an averaged error of 4% and the filler's sound speed with 1.7% of an average error. The material identification approach shows 67.3% success rate for the shell material identification and 80% for the filler identification.

The approach of material identification of a spherical targets shows reliable results for both synthetic and experiments datasets. The solution is demonstrated for spherical objects, but it can be extended for the cylinder shape as well. Further details are described in Chapter 6.

Chapter 6

Discussion and Conclusions

"Begin at the beginning," the King said, very gravely, "and go on till you come to the end: then stop."

Jane Carruth. Alice in Wonderland

6.1 Summary

In this thesis, subject of underwater object characterisation based on the processing of the wideband acoustic echoes was explored. Solutions for the material based classification and material recognition were developed and tested on synthetic and experimental data.

Chapter 2 presented a background review on the subjects involved into the research. It covered bio-inspired echolocation including research on the biosonar of dolphins and whales and bio-inspired pulses, adaptive echolocation, sound scattering from an object, understanding of the components of the echo and its relation to the object's properties, signal representation and overview of the machine learning techniques. This chapter looked at a number of techniques and considered their relative merit for the conducted research.

Chapter 3 described a solution for the numerical calculation of the sound scattering from a spherical shell immersed in water. The calculations were extended for the case of the 2-layer sphere with a liquid filler and metal shell from the work of Goodman and Stern [8] and Pailhas [9]. The results of the calculations were compared with experimental recordings on an example of an aluminium sphere, radius $0.075m$ and shell thickness $0.001m$, filled with fresh water. It had been demonstrated that the numerically calculated sound scattering is simulated very close to the real echo scattering and this approach can be used for data simulation. Using the numerical calculations, the chapter demonstrated influence of the sphere properties into the composition of the reflected pulses. The impact of the sphere's characteristics on the timing of the reflection components, their composition and positioning of the peak and notches in the echo's spectra had been highlighted in the chapter. This chapter presented a basis for the approaches which are introduced in Chapter 4 and Chapter 5.

Chapter 4 provided a solution for the machine learning based classification of the spheres. Machine learning techniques were applied and compared to distinguish between objects with different physical properties. Firstly, a multi-class classifier was applied to the echoes from a set of spherical shells. In this case, one class was presented by one sphere, which can be considered as a combination of characteristics, such as radius, thickness, shell and filler materials. The echoes were represented in time-frequency domain, using short time Fourier transform with different window lengths, and was fed into different classifiers, including convolution neural network, multilayer perceptron classifier, support vector machine, and gradient boosting. All of these approaches had previously been introduced and discussed in Chapter 2. The study of the results for different window lengths allowed analysis of the influence of time and frequency resolution on classification. The CNN provided the results with highest accuracy of $(98.44 \pm 0.8)\%$ over five object classes trained on grayscale TFD images with 0.1 ms window length of STFT. It was concluded, that a high performance of the classifier can be achieved with a window length, which provides fine trade off the time and frequency resolutions. This study presented good results for the classification of objects, which have the same shape and different physical properties. In the next step toward the material based classification, the objects were classified based on only one characteristic - filler material. Two classes of objects were formed for the study: water filled and air filled spherical shells. The objects were described by their form function, which was calculated from the recorded wideband echoes. The highest accuracy of $(98.8 \pm 1.31)\%$ was achieved with the form function descriptor and MLP classifier. The study demonstrated possibility of object classification based on a single characteristic of the objects - filler material. The presented machine learning approaches had a limited application, as objects with materials which have similar speeds of sound (for example water, salt water and oil) cannot be classified with high accuracy using these approaches. It led to the development of the method presented in Chapter 5.

Chapter 4 also presented a solution for a sea floor recognition using wideband pulses and machine learning techniques. The classification was based on the acoustic response from the seabed. Comparison of three different descriptors (time-frequency moment singular value decomposition (TFM-SVD), energy vector and complete frequency spectrum) were presented in the work. TFM-SVD descriptor provided the highest accuracy of 91.0% and showed the best result for the sea floor recognition task among the presented approaches.

In Chapter 5 a novel approach for the shell and filler materials identification was presented. The approach is based on the timing of the reflection components and doesn't require any training. The timing of the echoes components was calculated using matched filtering and a peak detector. A basic adaptive control of the minimum peak prominence for the peak detector was integrated into the solution. The approach is limited by a metal spherical shell filled with liquid with up to 2% thickness and tested for the diameter up to $0.3m$. The results for the material recognition methods were presented for synthetic and experimental data. The approach demonstrated 83.5% success rate for the shell material and 79.9% success rate for the filler material identification with the synthetic data, and 67.3%

success rate for the shell material identification and 80% for the filler identification with the experimental data. The approach of material identification provided reliable results for both synthetic and experiments data sets.

6.2 Future work

Cylindrical shell

The presented research is limited by the type of the objects, where only 2-layer spherical shell objects are considered. The solutions can be extended to cylindrical shells, which would allow to use the application for the underwater pipe and cable inspection. If the cylindrical object is insonified at a perpendicular angle, the material identification approach can be applied, as the trajectories of the plate wave and the wave reflected from the back wall are similar to the sphere. The machine learning based approaches have to be provided with a new training dataset for the cylindrical objects. It has to be considered, that switching to a shape which isn't axisymmetrical will introduce an angular dependence.

Mid-frequency enhancement

In Chapter 5, the material identification approach was build based on the timing on the two reflection components: the plate wave and the wave which propagates through the shell, the filler and reflects from the back wall of the sphere. These components provide sufficient information for the material identification. More sphere's parameters can be identified if another reflection component will be considered - mid-frequency enhancement (MFE). MFE contains information about the spherical shell characteristics. Zhang [59] and Kaduchak and Marson [60] were studying the enhancement, while Li *et al.* [41] illustrated how frequency and time delay of the mid-frequency enhancement can be used for the radius and thickness evaluation. The challenge of the MFE detection is that it can be overlapped in time by other components of the reflection [61]. Potentially, an approach for the automatic detection of the MFE can be developed. The extracted MFE parameters (frequency and time delay) can be used for the radius and thickness evaluation.

Adaptive sensing

It was illustrated in Chapter 5, that the bandwidth of the initial pulse influences the detection of the echo's components. Narrowing the pulse bandwidth led to the improvement, when there was a doubling peak problem. Nevertheless the solution has to be applied selectively, as it influences the pulse detector and demotes the filler and shell material identifier results for other cases. The bandwidth of the pulse could be adapted based on the feedback from the identifier, which processed the previous pulse. Concept of the approach is described in Chapter 2, Section 2.2. In this way, the adaptive echolocation can be implemented for

the case of the material identification approach. Potentially, this adaptive approach can be extended to the pulse duration and other pulse's parameters.

Multi-chirp pulses

In this work, single-chirp pulses are used for tasks of object classification and material identification. The single chirp pulses comply with the necessary requirement - provide wideband frequency range. The efficiency of the multi-chirp pulses is proven by nature, as it is used in echolocation by dolphins, whales, bats. Also double-chirp pulses were proposed and used by a number of authors. The complex multi-chirp structure of the pulses can be beneficial for the object classification and material identification tasks presented in this work.

6.3 Conclusion

This work presented an object characterisation in underwater environment using wideband sonar pulses. The object types were limited by a 2-layer spherical shell and numerical calculation of the sound scattering from the sphere is presented in this work. The research was focused on the material recognition for the better object characterisation. There are two different solutions presented in this work: a machine learning based classification approach and a timing based material identification approach. The machine learning based approach requires a training data and allows object classification based on their physical properties or their filler material. The novel timing based approach identifies material of the sphere's shell and filler from timing of the reflection components. It doesn't require any training data and provides an efficient solution for the material identification. Also in this thesis, the topic of adaptive echolocation was presented. Nevertheless, an integration of the adaptive approach to the presented object characterisation methods was assigned for the future work. This research presented novel and effective solutions for the task of underwater object characterisation using wideband sonar pulses.

References

[1] F. L. Chevalier, *Principles of Radar and Sonar Signal Processing*. Artech House, Inc, 2002.

[2] W. Au, R. Floyd, R. Penner, and A. E. Murchison, "Measurement of echolocation signals of the atlantic bottlenose dolphin, tursiops truncatus montagu, in open water," *Journal of the Acoustical Society of America*, vol. 56, 1974.

[3] W. Au, "The sonar of dolphins," *Acoustic Australia*, vol. 32, 2004.

[4] G. C. Gaunaurd, D. Brill, H. Huangb, P. W. B. Moore, and H. C. Strifors, "Signal processing of the echo signatures returned by submerged shells insonified by dolphin clicks: active classification," *Journal of the Acoustical Society of America*, vol. 103(3), 1998.

[5] C. Capus, Y. Pailhas, K. Brown, D. M. Lane, D. Houser, and P. Moor, "Bio-inspired wideband sonar signals based on observations of the bottlenose dolphin (tursiops truncatus)," *Journal of the Acoustical Society of America*, vol. 121(1), 2007.

[6] T. Leighton and P. White, "Dolphin-inspired target detection for sonar and radar," *Archives of Acoustics*, 2014.

[7] J. A. Fornshell and A. Tesei, "The development of sonar as a tool in marine biological research in the twentieth century," *International Journal of Oceanography*, 2013.

[8] R. R. Goodman and R. Stern, "Reflection and transmission of sound by elastic spherical shells," *Journal of the Acoustical Society of America*, vol. 34(3), 1962.

[9] Y. Pailhas, C. Capus, K. Brown, and P. Moor, "Identifying content of low profile target in cluttered environment using biosonar," *Underwater Acoustics conference, At Rhodes, Greece*, 2012.

[10] W. Au and J. Simmons, "Echolocation in dolphins and bats," *Physics Today*, vol. 60, 2007.

[11] W. W. Au, "History of dolphin biosonar research," *Acoustics Today*, vol. 11(4), 2015.

[12] L. Spallanzani, "Lettere sopra il sospetto di un nuovo senso nei pipistrelli (letters on the suspicion of a new sense in bats)," *Stamperia Reale (Royal Press). Turin, Italy*, 1794.

[13] G. Galati, *100 Years of Radar*. Springer, 2015.

[14] G. S. Pierce and D. R. Griffin, "Experimental determination of supersonic notes emitted by bats," *Journal of Mammology*, vol. 19, 1938.

[15] A. E. McBride, "Evidence for echolocation by cetaceans," *Deep Sea Reseach*, vol. 3, 1947.

[16] K. S. Norris, J. H. Prescott, P. V. Asa-Dorian, and P. Perkins, "An experimental demonstration of echo-location behavior in the porpoise, *Tursiops truncatus (Montagu)*," *Biological Bulletin*, vol. 120, 1961.

[17] S. H. Ridgway, *The Dolphin Doctor: A Pioneering Veterinarian Remembers the Extraordinary Dolphin That Inspired His Career*. Yankee Books, Dublin, NH, 1987.

[18] C. S. Johnson, "Auditory thresholds of the bottlenosed porpoise (*Tursiops truncatus, Montagu)*," *US Naval Ordnance Test Station, China Lake, CA*, 1966.

[19] F. G. Wood and W. E. Evans, "Adaptiveness and ecology of echolocation in toothed whales," *Animal Sonar Systems*, 1980.

[20] J. R. Potter and E. A. Taylor, "On novel reception models for bottlenose dolphin echolocation," *Proceedings of the Institute of Acoustics*, vol. 24, 2001.

[21] R. A. Kastelein, A. Y. Supin, and J. A. Thomas, eds., *Marine Mammal Sensory Systems*. Springer, 1992.

[22] J. E. Sigurdson, "Analysing the dynamics of dolphin biosonar behavior during search and detection tasks," *Proceedings of the Institute of Acoustics*, vol. 19, no. 9, 1997.

[23] J. T. Fulton, "Dolphin biosonar echolocation. a case study," *Process in Biological Hearing. Corona Dek Mar, Ca. USA*, 2005.

[24] M. Wahlberg, K. Beedholm, A. Heerfordt, and B. Mohl, "Demonstration of adaptation in beluga whale echolocation signals," *Journal of the Acoustical Society of America*, vol. 130(5), 2011.

[25] D. Herzing and M. E. dos Santos, *Functional aspects of echolocation in dolphins*. Chicago, USA: University of Chicago Press, 2004.

[26] S. D. Ibsen, W. Au, P. E. Nachtigall, and M. Breese, "Functional bandwidth of an echolocating atlantic bottlenose dolphin (*Tursiops truncatus*)," *Journal of the Acoustical Society of America*, vol. 125, 2009.

[27] W. Au, D. Carder, R. Penner, and B. Scronce, "Demonstration of adaptation in beluga whale echolocation signals," *Journal of the Acoustical Society of America*, vol. 77(2), 1985.

[28] D. Houser, D. Helweg, and P. Moor, "Classification of dolphin echolocation clicks by energy and frequency distributions," *Journal of the Acoustical Society of America*, vol. 106, 1999.

[29] M. W. Muller, J. S. Allen, W. W. L. Au, and P. E. Nachtigal, "Time-frequency analysis and modeling of the backscatter of categorized dolphin echolocation clicks for target discrimination," *Journal of the Acoustical Society of America*, vol. 124, no. 1, 2008.

[30] T. G. Leighton, G. H. Chua, and P. R. White, "Do dolphins benefit from nonlinear mathematics when processing their sonar returns?," *Proceedings of the Royal Society*, vol. Published online, 2012.

[31] T. G. Leighton, D. C. Finfer, G. H. Chua, P. R. White, and J. K. Dix, "Clutter suppression and classification using twin inverted pulse sonar in ship wakes," *Acoustical Society of America*, vol. 130, 2011.

[32] Y. Pailhas, C. Capus, K. Brown, and P. Moor, "Analysis and classification of broadband echoes using bio-inspired dolphin pulses," *Journal of the Acoustical Society of America*, vol. 127, 2010.

[33] G. Qiao, X. Qing, D. Nie, Y. Zhang, and J. Tanh, "Underwater cylindrical shell in different thickness recognition using biomimetic dolphin clicks," *Proceedings of Ocean'16*, 2016.

[34] M. Dmitrieva, K. Brown, and D. Lane, "Adaptive bio-inspired signals for better object characterisation," *Proccidings of the Living Machine conference*, 2015.

[35] M. Dmitrieva, K. Brown, G. Heald, and D. Lane, "Image processing based strategy for adaptive wideband echolocation," *Student Forum. ICASSP, New Orleans*, 2017.

[36] J. W. S. Rayleigh, *Theory of sound*. London: Macmillan and co, 1877.

[37] P. M. Morse, *Vibration and Sound*. New York: McGraw-Hill Book Company, 1936.

[38] A. Lowan, P. M. Morse, H. Freshbach, and M. Lax, *Scattering and Radiation from Circular cylinders and spheres*. Wachington, D.C.: U.S. Navy Department, Office of Research and Inventions, 1946.

[39] V. C. Anderson, "Sound scattering from a fluid sphere," *The Journal of the Acoustical Society of America*, vol. 22, 1950.

[40] M. C. Junger, "Radiation loading of cylindrical and spherical surfaces," *Journal of the Acoustical Society of America*, vol. 24, 1952.

[41] W. Li, G. R. Liu, and V. K. Varadan, "Estimation of radius and thickness of a thin spherical shell in water using the midfrequency enhancement of a short tone burst response," *Journal of the Acoustical Society of America*, vol. 118(4), 2005.

[42] G. S. Sammelmann, D. H. Trivett, and R. H. Hackman, "The acoustic scattering by a submerged, spherical shell. I: The bifurcation of the dispersion curve for the spherical antisymmetric lamb wave," *Journal of the Acoustical Society of America*, vol. 85, 1989.

[43] J. J. Faran, "Sound scattering by solid cylinders and spheres," *Journal of the Acoustical Society of America*, vol. 23(4), 1951.

[44] A. Love, *A treatise on the mathematical theory of elasticity*. Cambridge: at the university press, 1892.

[45] R. Hickling, "Analysis of echoes from a solid elastic sphere in water," *Acoustical Society of America*, vol. 34, 1962.

[46] R. Doolittle and H. Uberall, "Sound scattering by elastic cylindrical shells," *Acoustical Society of America*, vol. 39, 1965.

[47] Y. Pailhas, C. Capus, K. Brown, and Y. Petillot, "Design of artificial landmarks for underwater simultaneous localisation and mapping," *IET Radar Sonar Navig.*, 2013.

[48] P. A. Chinnery and V. F. Humphrey, "Experimental visualization of acoustic resonances within a stadium-shaped cavity," *Physical Review Letters of the American Physical Society*, 1996.

[49] P. A. Chinnery, J. Zhang, and V. F. Humphrey, "Acoustic scattering by nonmetallic and metallic cubes in the elastic resonance regime: Experimental measurements and combined finite element/boundary element modeling," *Journal of the Acoustical Society of America*, vol. 102(1), 1997.

[50] P. A. Chinnery, V. F. Humphrey, and J. Zhang, "Low-frequency acoustic scattering by a cube: Experimental measurements and theoretical predictions," *Journal of the Acoustical Society of America*, vol. 101(5), 1997.

[51] L. L. Thompson, "A review of finite-element methods for time-harmonic acoustics," *Journal of the Acoustical Society of America*, vol. 119, 2006.

[52] S. Kirkup, *The Boundary Element Method in Acoustics*. Journal of Computational Acoustics, 2007.

[53] A. Dariouchy, E. Aassif, G. Maze, D. Decultot, and A. M. and, "Prediction of the acoustic form function by neural network techniques for immersed tubes," *Journal of the Acoustical Society of America*, vol. 124, 2008.

[54] R. Hickling, *Acoustic Radiation and Reflection from spheres*. California Insitute of Technology, 1963.

[55] S. G. Kargl and P. L. Marston, "Observations and modeling of the backscattering of short tone bursts from a spherical shell: Lamb wave echoes, glory, and axial reverberations," *Journal of the Acoustical Society of America*, vol. 85(3), 1989.

[56] H. Lamb, "On wave in an elastic plate," *Proceesings of the royal society*, vol. 93, pp. 114–128, 1917.

[57] G. S. Sammelmann and R. H. Hackman, "The acoustic scattering by a submerged, spherical shell. II: The high-frequency region and the thickness quasiresonance," *Journal of the Acoustical Society of America*, vol. 89, 1991.

[58] G. S. Sammelmann and R. H. Hackman, "The acoustic scattering by a submerged, spherical shell. III: Pole trajectories in the complex-ka plane," *Journal of the Acoustical Society of America*, vol. 90, 1991.

[59] L. Zhang, N. Sun, and P. Marston, "Midfrequency enhancement of the backscattering of tone bursts by thin spherical shells," *Journal of the Acoustical Society of America*, vol. 91, 1991.

[60] G. Kaduchak and P. L. Marston, "Backscattering of chirped bursts by a thin spherical shell near the coincidence frequency," *Journal of the Acoustical Society of America*, vol. 93, 1993.

[61] S. D. Anderson, *Space-Time-Frequency processing from the Analysis of Bistatic Scaterring for Simple Underwater Targets*. PhD Thesis Georgia Institute of Technology, 2012.

[62] G. Qiao, X. Qing, and D. Nie, "Acoustic scattering from stainless steel shells with varying wall thickness using a biomimetic click: Modeling and interpretation," *Journal of the Acoustical Society of America. Published online*, 2017.

[63] A. Tesei, W. L. J. Fox, A. Maguer, and A. Lovikc, "Target parameter estimation using resonance scattering analysis applied to air-filled, cylindrical shells in water," *Journal of the Acoustical Society of America*, vol. 108, 2000.

[64] B. Boashash, *Time-Frequency Signal Analysis and Processing*. Elsevier, 2016.

[65] A. V. Oppenheim, R. W. Schafer, and J. R. Buck, *Discrete-Time Signal Processing*. NJ, USA: University of Chicago Press, 1999.

[66] A. V. Oppenheim, A. S. Willsky, and S. H. Nawab, *Signals and systems (2nd ed.)*. NJ, USA: University of Chicago Press, 1996.

[67] P. R. White, "The characterisation of bio-acoustic signals through time-frequency methods," *In Proceedings of the Institute of Acoustics 2nd Symposium on Underwater Bio-Sonar and Bioacoustic Systems. Institute of Acoustics*, 2001.

[68] J. Starkhammar and M. Hansson-Sandsten, "Evaluation of seven time-frequency representation algorithms applied to broadband echolocation signals," *Advances in Acoustics and Vibration*, vol. 56, 2015.

[69] D. Gabor, "Theory of communication," *Journal of the Institution of Electrical Engineers*, vol. 93, 1946.

[70] J. Ville, "Theorie at applications de la notion de signal analytique," *Cables et Transmission*, vol. 2A, 1948.

[71] L. Cohen, "Time-frequency distirbution - a review," *Proceesings of the IEEE*, 1989.

[72] J. Allen and L. Rabiner, "A unified approach to short-time fourier analysis and synthesis," *Proceesings of the IEEE*, vol. 65, 1977.

[73] L. Cohen, *Time-frequency analysis: theory and applications*. New York, USA: Printice-Hall, 1995.

[74] H.-I. Choi and W. J. Williams, "Improved time-frequency representation of multicomponent signals using exponential kernels," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 6, 1989.

[75] M. Hansson-Sandsten, "Optimal multitaper wigner spectrum estimation of a class of locally stationary processes using hermite functions," *EURASIP Journal on Advances in Signal Processing*, 2011.

[76] P. Loughlin, J. Pitton, and B. Hannaford, "Approximating time-frequency density functions via optimal combinations of spectrograms," *IEEE Signal Processing Letters*, vol. 1, no. 12, 1994.

[77] E. P. Wigner, "On the quantum correction for thermodynamic equilibrium," *American Physical Society*, vol. 40, no. 5, 1932.

[78] P. Flandrin, F. Auger, and E. Chassande-Mottin, "Time-frequency reassignment: from principles to algorithms," in *Applications in Time-Frequency Signal Processing* (A. Papandreou-Suppappola, ed.), CRC Press, 2003.

[79] F. Auger and P. Flandrin, "Improving the readability of time-frequency and time-scale representations by the reassignment method," *IEEE Transactions of signal processing*, vol. 43, 1995.

[80] Y. S. Shin and J.-J. Jeon, "Pseudo wigner-ville time-frequency distribution and its application to machinery condition monitoring," *Shock and Vibration*, vol. 1, 1993.

[81] S. Marsland, *Machine Learning. An Algorithmic Perspective*. CRC Press, 2015.

[82] S. B. Kotsiantis, "Supervised machine learning: A review of classification techniques," *Emerging Artificial Intelligence Applications in Computer Engineering*, 2007.

[83] L. Guerra, L. McGarry, V. Robles, C. Bielza, P. Larranaga, and R. Yuste, "Comparison between supervised and unsupervised classifications of neuronal cell types: A case study," *Developmental Neurobiology*, vol. 71(1), 2011.

[84] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A training algorithm for optimal margin classifiers," in *Proceedings of the 5th Annual ACM Workshop on Computational Learning Theory*, ACM Press, 1992.

[85] V. Vapnik, "An overview of statistical learning theory," *IEEE Transactions on Neural Networks*, vol. 10, Sep 1999.

[86] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20(3), Sep 1995.

[87] V. R. Jakkula, "Tutorial on support vector machine (svm)," Washington State University, 2011.

[88] J. M. Moguerza and A. Muoz, "Support vector machines with applications," *Statistical Science*, vol. 21, no. 3, 2006.

[89] J. H. Friedman, "Greedy function approximation: A gradient boosting machine," *The Annals of Statistics*, vol. 29, no. 5, 2001.

[90] J. H. Friedman, "Stochastic gradient boosting," *Computational Statistics and Data Analysis*, vol. 38, no. 4, 2002. Nonlinear Methods and Data Mining.

[91] P. Bhlmann and T. Hothorn, "Boosting algorithms: Regularization, prediction and model fitting," *Statistical Science*, vol. 22, no. 4, 2007.

[92] W. McCulloch and W. Pittsins, "A logical calculus of the ideas immanent in nervous activity," *Bulletin of Mathematical Biophysics*, vol. 5, 1943.

[93] M. T. Hagan, H. B. Demuth, M. H. Beale, and O. D. Jesus, *Neural network Design*. eBook, 2014.

[94] S. Shalev-Shwartz and S. Ben-David, *Understanding Machine Learning: From Theory to Algorithms*. Cambridge University Press, 2014.

[95] R. Hecht-Nielsen, "Theory of the backpropagation neural network," in *International 1989 Joint Conference on Neural Networks*, vol. 1, 1989.

[96] J. A. Bullinaria, "Lecture notes in neural computation. Lecture: Applcation of multilayer perceptron," 2015.

[97] M. J. Crocker, *Handbook of Acoustics*. John Wiley and Sons, 1998.

[98] M. Dmitrieva, M. Valdenegro-Toro, K. Brown, G. Heald, and D. Lane, "Object classification with convolution neural network based on the time-frequency representation of their echo," *Proceedings of the MLSP workshop, Tokyo*, 2017.

[99] M. Dmitrieva, K. Brown, G. Heald, and D. Lane, "Classification of spherical objects based on the form function of acoustic echoes," *Proceedings of the UACE 2017, Skiathos*, 2017.

[100] M. Dmitrieva, K. Brown, G. Heald, and D. Lane, "Sea floor recognition with bio-inspired echolocation," *Proceeding of the OCEANS, Aberdeen, UK*, 2017.

[101] N. Valeyrie, Y. Pailhas, C. Capus, and S. Dugelay, "ONMEX and MANEX 2016 CMRE sea trials with a wideband multi-beam sonar system," *Proceedings of the UACE 2017, Skiathos*, 2017.

[102] G. S. Kino, *Acoustic waves: devices, imaging, and analog signal processing*. Prentice-Hall, 1987.

[103] F. J. Harris, "On the use of windows for harmonic analysis with the discrete fourier transform," *Proceedings of the IEEE*, vol. 66, 1978.

[104] R. O. Nielsen, *Sonar Signal Processing*. Artech House, 1991.

[105] D. Middleton, *An Introduction to Statistical Communication Theory*. McGraw-Hill, 1960.

[106] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *Journal of Machine Learning Research*, vol. 15(1), 2014.

[107] J.-P. Vert, K. Tsuda, and B. Schölkopf, *A Primer on Kernel Methods*. Cambridge, MA, USA: MIT Press, 2004.

[108] A. Akhbardeh, S. Junnila, M. Koivuluoma, T. Koivistoinen, and A. Varri, "Applying novel time-frequency moments singular value decomposition method and artificial neural networks for ballistocardiography," *EURASIP Journal on Advances in Signal Processing*, 2006.

[109] G. Turin, "An introduction to matched filters," *IRE Transactions on Information Theory*, vol. 6, 1960.